Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Long-term memory detection with bootstrapping techniques: empirical analysis

Author: **Branislav Albert**

Supervisor: **PhDr. Ladislav Krištoufek**

Academic Year: **2011/2012**

Declaration of Authorship

I hereby declare that I compiled this thesis on my own under the leadership of my supervisor, using only the listed resources and literature.

I grant the permission to Charles University to reproduce and distribute copies of this thesis document in whole or in part.

in Prague on July 31, 2012

Signature

Acknowledgments

I am especially grateful to my thesis adviser PhDr. Ladislav Krištoufek for useful comments and ideas.

A part of this thesis was written while the author was at the University of Konstanz.

Bibliographic entry

ALBERT, B. (2012): "Long-term memory detection with bootstrapping techniques: empirical analysis." *(Unpublished bachelor thesis).* Charles University in Prague. Supervisor: PhDr. Ladislav Krištoufek.

Length: 15,579 words.

Abstract

A time series has long range dependence if its autocorrelation function is not absolutely convergent. Presence of long memory in a time series has important consequences for consistency of several time series estimators and forecasting. We present a self-contained theoretical treatment of time series models necessary for study of long range dependence and survey a large list of parametric and semiparametric estimators of long range dependence. In a Monte Carlo study, we compare size and power properties of four estimators, namely R/S, DFA, GPH and Wavelet based method, when relying on asymptotic normality of the estimators and distributions obtained from the moving block bootstrap. We find out that the moving block bootstrap can improve the size of the R/S estimator. In general however, the moving block bootstrap did not perform satisfactorily for other estimators, while GPH and Wavelet estimators offer the most reliable asymptotic confidence intervals.

Keywords bootstrapping, moving block bootstrap, long-term memory, time series, R

Author's e-mail `branoalbert@gmail.com`
Supervisor's e-mail `kristoufek@ies-prague.org`

Abstrakt

Časová rada má dlhú pamäť ak jej autokorelačná funkcia nie je absolútne konvergentná. Prítomnosť dlhej pamäte v časovej rade má dôležité následky pre konzistentnosť niekoľkých estimátorov z oblasti časových rad a pre predpovedanie. V tejto práci prezentujeme ucelený prehľad modelov časových rad nevyhnutných pre štúdium dlhej pamäte a následne sa zameriavame na množstvo parametrických a semiparametrických estimátorov dlhej pamäte. V

Monte Carlo štúdii porovnávame pravdepodobnosť chyby prvého typu a silu štyroch estimátorov, menovite R/S, DFA, GPH a metóde založenej na Waveletoch, pre asymptoticky normálne rozdelenie estimátorov a rozdelenia získané pomocou metódy moving block bootstrap. Zisťujeme, že moving block bootstrap dokáže zlepšiť pravdepodobnosť chyby prvého typu u estimátora R/S. Vo všeobecnosti však moving block bootstrap neprináša uspokojivé výsledky. Estimátory GPH a Wavelet ponúkajú najspoľahlivejšie asymptotické intervaly spoľahlivosti.

Klíčová slova	bootstrapping, moving block bootstrap, dlhá pamäť, časové rady, R

E-mail autora	`branoalbert@gmail.com`
E-mail vedoucího práce	`kristoufek@ies-prague.org`

Contents

List of Tables

List of Figures

Acronyms

LRD Long Memory or Long Range Dependence

MC Monte Carlo

MBB Moving Block Bootstrap

FIP Fractionally Integrated Processes

SAX Slovak Share Index

WN White Noise

MA Moving Average

ACF Autocorrelation Function

ARMA Autoregressive Moving Average

ARFIMA Autoregressive Fractionally Integrated Moving Average

R/S Rescaled Range

DFA Detrended Fluctuation Analysis

GPH Geweke and Porter-Hudak

Chapter 1

Introduction

Although the notion of processes with Long Memory or Long Range Dependence (LRD) could be traced back to the 50s and the work of British hydrologist Harold Edwin Hurst, especially in (Hurst, 1951), economics was relatively slow to reflect on the importance of these phenomena (Baillie, 1996). Thus the assumption of no persistence in the autocorrelations of time series underlies a major portion of econometric theory. Empirically driven relaxation of this assumption, however, has led to applications in geophysics and hydrology (seismology, wind speed, climate effects), medicine (blood pressure), technology (highway and internet traffic) and many others (Kantelhardt, 2009). The results offer important recommendations for system designs and modelling, congestion and flow control, reliable predictions and simulations (Hernandez-Campos *et al.*, 2011).

In economics, traditional theory of asset pricing and its assumptions disregarded the possibility of long-range dependence both in the market time series themselves and also in several important transformations of the series. Current research, on the other hand, suggests that presence of LRD in the market time series has important implications for diverse areas from macroeconomics to risk management, especially in regard to responses to unanticipated shocks, volatility modelling and forecasting (Taylor, 2000) and (Henry & Zaffaroni, 2002). This makes the processes with long memory interesting both scientifically and practically.

Current scientific literature provides various statistical tests for long range dependence which differ in several aspects, notably efficiency and assumptions about the underlying data generating processes. While the tests focus on estimating either the Hurst parameter H or fractional difference parameter d, extraction of confidence intervals and thereby statistical inference for these estimators is a more complicated but similarly important task. Moreover, Teverovsky *et al.* (1999) states that exact distributions of these estimators are often not known or are known only asymptotically with arguably imprecise finite-sample approximation.

The major objective of this paper is to compare the quality of asymptotic confidence intervals of four long range dependence estimators (i.e. R/S, DFA, GPH, and Wavelet-based method) with confidence intervals obtained by performing Moving Block Bootstrap (MBB) in a Monte Carlo study by comparing their respective size and power properties. MBB is a modification of the original bootstrap proposed by Efron (1982) used in time series framework and it can provide approximate distribution of a time series statistic by randomly resampling blocks from the original time series (Kuensch, 1989).

Other parts of this paper seek to develop a self-contained theoretical treatment of the time series processes with LRD and apply the long memory tests to the assessment of Slovak Share Index (SAX).

The thesis is structured as follows: Chapter 2 discusses stochastic processes in general and we focus on the tests for LRD in Chapter 3. Chapter 4 presents Monte Carlo (MC) study. In Chapter 5 is an application to real-world time series. Chapter 6 presents the summary our findings.

Chapter 2

Time Series Statistics

This chapter develops necessary theoretical treatment of many important time series models which is a necessary prerequisite for any further discussions. After presenting basic definitions, we proceed with a bottom-up approach in the ARFIMA framework. We focus on models of volatility in the third part to present a self-contained and thorough analysis of the contemporary time series models needed for the analysis of long memory.

2.1 Basic definitions

A stochastic process is a collection of random variables that evolves over time. We denote a stochastic process as $\{y_t\}$, or simply y_t, and the random variable in time period t as Y_t. An ordered sequence of observations from a stochastic process is called a time series. The notion of times series is closely connected to the theory of stochastic processes because we understand any time series as a single realization of some underlying stochastic process. A significant portion of the theory of time series is devoted to the determination of a possible underlying stochastic process from some observed time series.

Each observation from a time series is a realization of a single random variable from the collection of random variables in the corresponding stochastic process. In time series analysis, we can usually obtain just a single sample of any particular stochastic process so that we have precisely one observation for any moment in time. We usually imagine a stochastic process as a one stretching infinitely into the past and future, that is $\{y_t\}_{t=-\infty}^{+\infty}$. Our observed sample, $\{y_t\}_{t=1}^{T}$, is thus necessarily a subset of this theoretical process. We will denote the period of observation as τ, that is, $\tau = 1, \cdots, T$.

To simplify notation in the following text, it is handy to introduce a lag operator L (sometimes also referred to as a backshift operator, B). Applying a lag operator on a time series y_t creates a new time series, denoted y_{t-1}, for which the value at time t equals the value that the original time series acquired at time $t-1$. We denote this operation as $Ly_t = y_{t-1}$. Lag operator follows similar rules as that of multiplication, especially commutative, associative and distributive laws (Shumway & Stoffer, 2006).

We will focus our discussions on the theory of stationary stochastic processes. A stationary stochastic process $\{y_t\}$ (sometimes also called strictly stationary) is a process for which joint probability distribution of any set of its random variables is time invariant. Following Wooldridge (2009):

Definition 2.1 (Stationary stochastic process). A stochastic process $\{y_t\}$ is stationary if for any collection of its individual random variables at times $1 \leq t_1 \leq \ldots \leq t_n \leq T$, the joint distribution function of $(y_{t_1}, \ldots, y_{tn})$ is the same as that of $(y_{t_1+h}), \ldots, y_{t_n+h})$ for all integers $h \geq 1$.

For many practical considerations, the requirements put forward in the definition of strict stationarity may be too restrictive. We therefore focus primarily on covariance stationary processes.

Definition 2.2 (Covariance stationary stochastic process). A stochastic process with finite second moment ($\mathbb{E}(y^2) < \infty$) $\{y_t\}$ is covariance stationary (or weakly stationary) if:

- Expected value of y_t does not depend on the index t, that is $E(y_t) = \mu$ and
- Covariance of y_t and y_{t-i} does not depend on the index t, that is $Cov(y_t, y_{t-i}) = \gamma_i$ for $i = 0, 1, \ldots \infty$

It follows from the two preceding definitions that any stationary process that fulfils the condition of finite second moment is also a covariance stationary process - probability distribution of any particular random variable in stationary process is the same as well as joint probability distribution for any pair of two random variables which leads to the same expected value and covariance.

The reversed implication does not hold both because the time dependency of higher orders is allowed in covariance stationary time series and also because the definition of covariance stationary processes does not discuss the exact probability distributions but only two parameters of these distributions, Hamilton (1994) offers some intuitive examples for this distinction.

Correlation between y_t and y_{t-k} of a covariance stationary process will be called the kth autocorrelation, denoted ρ_k. For a stationary process, it holds from definition of correlation that: $\rho_k = \frac{\gamma_k}{\gamma_0}$, where γ_k is kth autocovariance of y_t. We can plot several autocorrelations with respect to the lag to obtain autocorrelation function. For $i = 0$, the autocorrelation is equal to one. Since this is always the case for any time series, including the lag number zero in the plot of autocorrelation function is not particularly meaningful.

A similar concept to that of autocorrelation is *partial* autocorrelation. hth partial autocorrelation of y_t, ϕ_{hh}, can be computed in terms of autocovariances of y_t as:

$$\begin{pmatrix} \phi_{1,h} \\ \phi_{2,h} \\ \vdots \\ \phi_{h,h} \end{pmatrix} = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{h-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{h-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{h-1} & \gamma_{h-2} & \cdots & \gamma_0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_h \end{pmatrix}$$

It holds that $\phi_{11} = \rho_1$ for any time series process.

A convenient estimator of the hth partial autocorrelation comes from performing an OLS regression of y_t on its h lags:

$$y_t = \hat{c} + \hat{\phi}_{1,h} y_{t-1} + \cdots + \hat{\phi}_{h,h} y_{t-h} + \varepsilon_t.$$

Current literature on LRD is abundant and so is the amount of various definitions of this phenomenon. But although the definitions differ in exact wordings or emphasis, they share the underlying notion of slowly decreasing autocorrelation function. This can be formalized as follows:

Definition 2.3 (Long Range Dependence). We say that a covariance stationary time series has long memory if the autocorrelation function is not absolutely convergent, i.e. $\sum_{i=1}^{\infty} |\rho_i|$ diverges.

Since autocorrelation function is not limited to positive or non-negative values, we impose a restriction of the absolute value of the function. Another authors, notably (Robinson *et al.*, 2003, ch. 1), work with original autocorrelations and it is important to stress that these definitions are not equivalent as simple convergence implies absolute convergence but not vice versa.

The autocorrelation function of a long memory time series is widely assumed to follow hyperbolic decay proportional to k^{2H-2}. This means that

$$\lim_{k\to\infty} \frac{\rho_k}{c \cdot k^{2H-2}} = 1.$$

Parameter H is the Hurst exponent or Hurst coefficient. We can see from the formula that the process will exhibit long memory for $H \in (1/2, 1)$ while with $H = 1/2$ is the process either independent or its autocorrelation function absolutely converges and it is dubbed as a process with short memory. For $H \in (0, 1/2)$ is the process negatively dependent or anti-persistent (Robinson & Henry, 1999).

We have defined LRD exclusively for covariance stationary series due to simplicity and clarity concerns. The overall notion would not be changed even if we allowed for general stationary processes.

An equivalent way to define long memory is in the spectral density framework since any covariance stationary process has both time and frequency domain representation. Following Hamilton (1994), population spectrum of y, $f(\lambda)$, can be derived from its autocovariance generating function g as

$$f(\lambda) = \frac{1}{2\pi} g(e^{-i\lambda}) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^i \lambda j,$$

given that the process does not have LRD. This expression simplifies to:

$$f(\lambda) = \frac{1}{2\pi} \left[\gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j cos(\lambda j) \right].$$

We can draw similar conclusions from the spectrum as we did from the Hurst parameter. If $f(0) = \infty$, i.e. the spectrum has a "pole" at frequency zero, the process has long memory. It is anti-persistent if $f(0) = 0$ and has short memory for $f(0) \in (0, \infty)$.

2.2 Processes in the ARFIMA form

2.2.1 White Noise

White Noise (WN) is probably the simplest stochastic process from the statistical perspective but it serves as a basic building block for several more complicated models. Its name is derived from acoustics, the term *noise* could be thought of as a random unwanted component of a variable of interest while the term *white* comes from white light which can be decomposed into full color spectrum, i.e. into a color spectrum without a dominant frequency.

Definition 2.4 (White Noise). A stochastic process $\{\varepsilon_t\}_{t=-\infty}^{+\infty}$ is called White Noise if:

- Expected value of each random variable ε_t is zero, $\mathbb{E}(\varepsilon_t) = 0$ and
- Variance of each random variable ε_t is constant across time, $Var(\varepsilon_t) = \mathbb{E}(\varepsilon_t^2) = \sigma^2$ and
- Each ε_i and ε_j are uncorrelated for $i \neq i$, $Corr(\varepsilon_i, \varepsilon_j) = \frac{\mathbb{E}(\varepsilon_i \cdot \varepsilon_j)}{\sigma^2} = 0$

If we replace the third condition with a little stronger requirement of independence, we arrive at so called *Independent White Noise*.

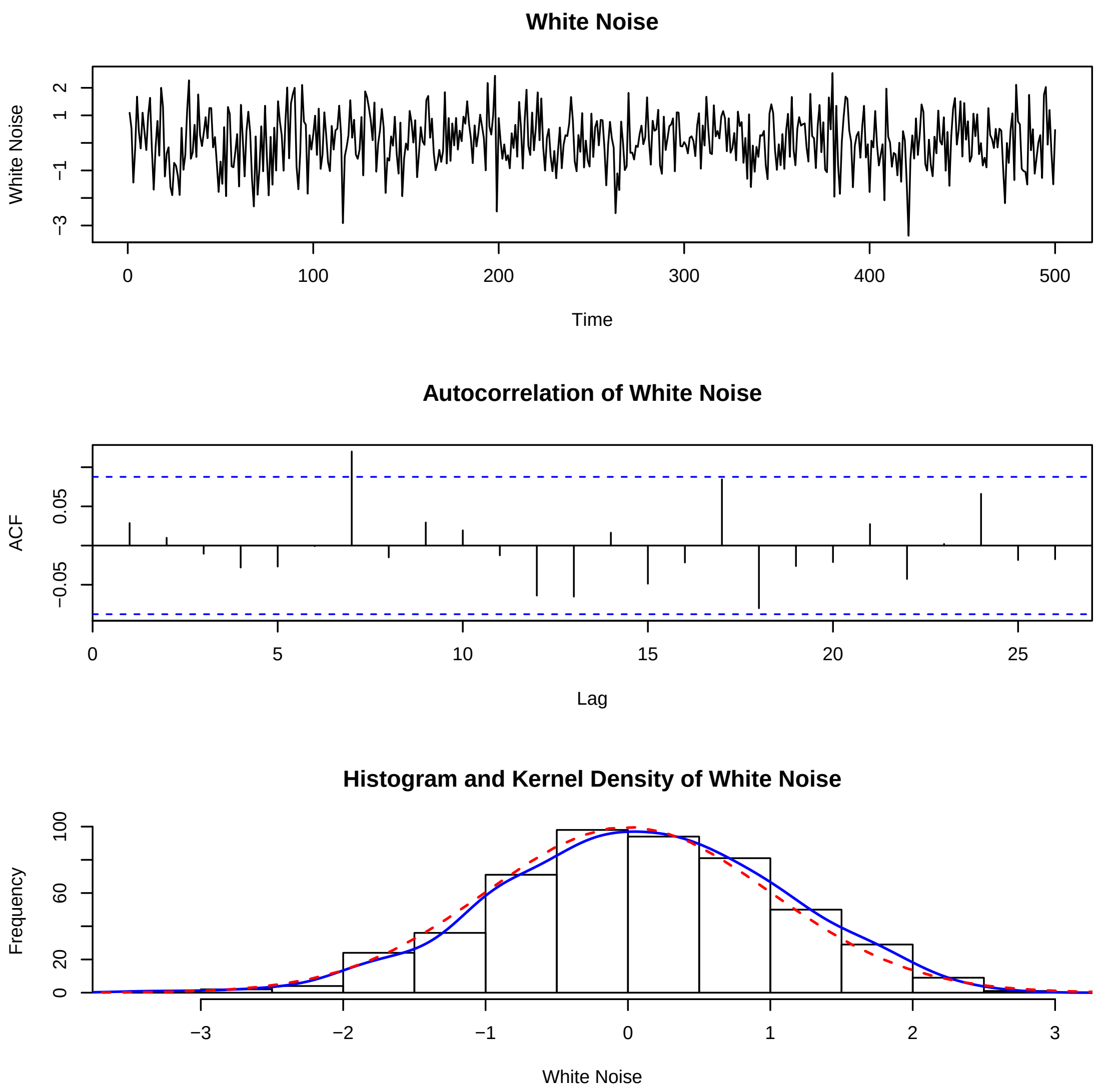
We can also specify a particular distribution for the individual random variables ε_t, the usual one being normal distribution. Therefore, if $\varepsilon_t \sim N(0, \sigma^2)$ we talk about *Gaussian Noise*.

First graph in Figure 2.1 is an example of a realization of the Gaussian White Noise with standard normal distribution of the individual ε_t. It is characterized by high randomness in the laymen meaning of the word, the plot does not contain any apparent regularities, trends or cycles.

The second graph is the graph of the autocorrelation function of this particular time series. Another interesting feature of this graph is that sample autocorrelations (especially for $i \geq 1$) are non-zero even though the underlying process assumes uncorrelated individual observations. It was however shown, see (Ding *et al.*, 1993, ch. 3), that sample individual autocorrelations of a White Noise follow $N(0, 1/T)$, $\hat{\rho}_i \sim N(0, 1/T)$. This leads to 95% confidence interval in form $\frac{\pm 1.96}{\sqrt{T}}$. In our case, $\frac{\pm 1.96}{\sqrt{T}} = \frac{\pm 1.96}{\sqrt{500}} \approx \pm 0.087$. 24 out of 25 autocorrelations are within this 95% confidence interval (96%) which is in good agreement with the theory.

Histogram and kernel density fit conveys a strong resemblance to the standard normal distribution, plotted with dashed line, and are included for control purposes. Note that kernel density graph and probability density function of the standard normal distribution are scaled by $0.5 \cdot 500$ to fit the histogram.

Figure 2.1: White Noise, $\varepsilon_t \sim N(0, 1)$



Source: author's computations.

2.2.2 Moving Average of order q

Moving Average (MA) processes are the most natural extension of the WN processes.

Definition 2.5 (Moving Average of order q). A stochastic process $\{y_t\}_{t=-\infty}^{+\infty}$ is said to be Moving Average process of order q if it satisfies:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \tag{2.1}$$

in which the $\{\varepsilon_t\}$ is a White Noise process from Definition 2.4. $\theta_1, \ldots, \theta_q \in \Re, \theta_q \neq 0, q \in \mathbb{N}$.

In this case, the properties of MA(q) are as follows (Hamilton, 1994, ch. 3.3):

- $\mathbb{E}(Y_t) = \mathbb{E}(\mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}) = \mu$
- $Var(Y_t) = \gamma_0 = \mathbb{E}(Y_t - \mu)^2 = \cdots = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2$
- $Cov(Y_t, Y_{t-j}) = \gamma_j = \begin{cases} (\theta_j + \theta_{j+1}\theta_1 + \cdots + \theta_q \theta_{q-j})\sigma^2 & for\ j = 1, 2, \ldots, q \\ 0 & for\ j > q. \end{cases}$

These properties are not hard to prove and they are based on the fact that individual random variables in the WN are uncorrelated. We can see that MA(q) fulfills the conditions from Definition 2.2 for a covariance stationary process, because its second moments do not depend on time and are finite. Moreover, MA(q) does not have long memory because q is finite.

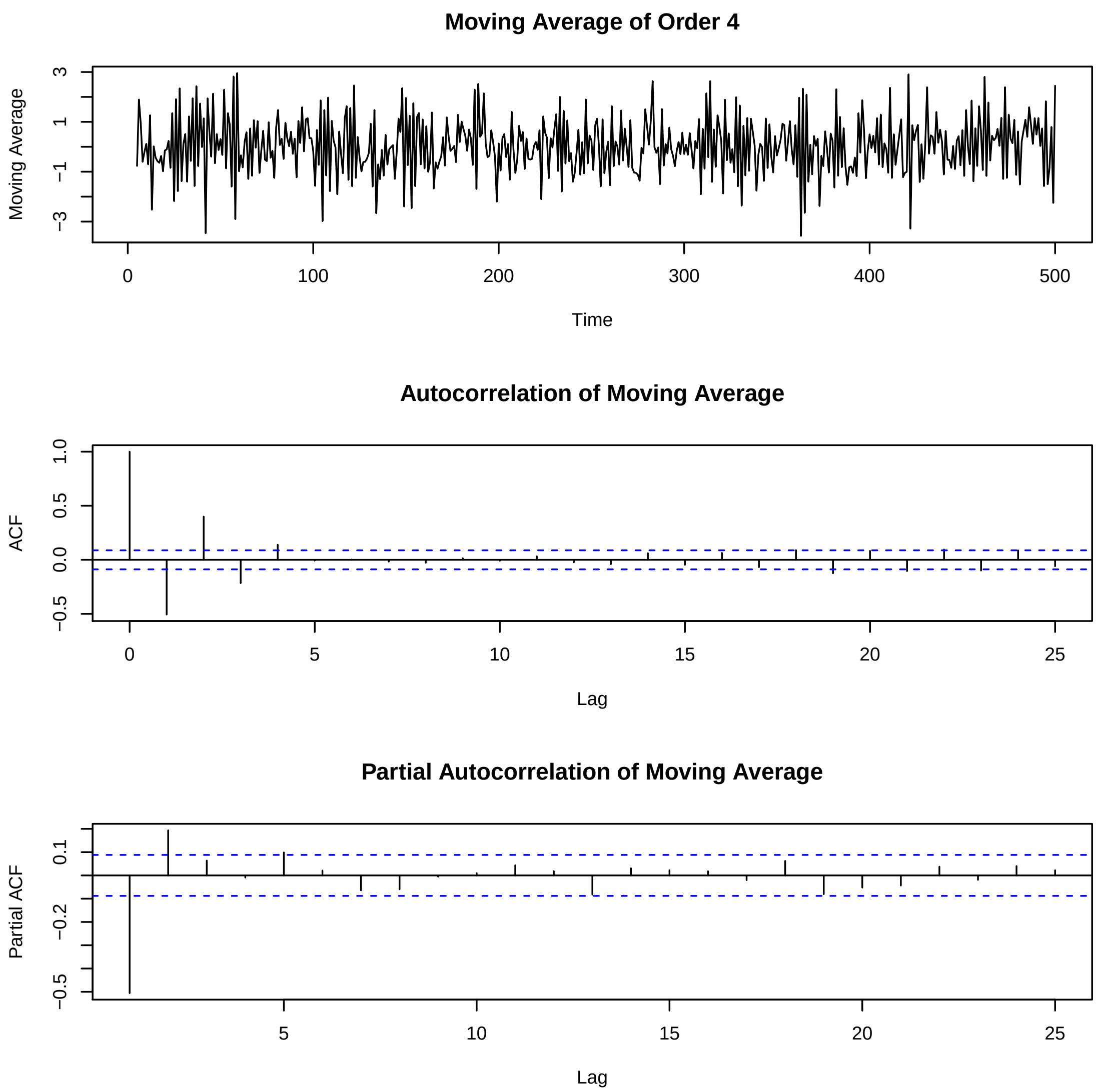
In theory, autocorrelation function of an MA(q) process will always have a cut-off point at lag q. Partial autocorrelation function, on the other hand, does not have a cut-off point but rather decreases to zero in limit only.

Figure 2.2 presents an example of MA(4) process, $Y_t = \varepsilon_t - 0.5\varepsilon_{t-1} + 0.4\varepsilon_{t-2} - 0.3\varepsilon_{t-3} + 0.2\varepsilon_{t-4}$, with underlying Gaussian Noise with standard deviation equal to unity. The example purposefully exhibits non-trivial autocorrelations for the first 4 lags while the following autocorrelations are around zero. We can also see that at lag zero, computed autocorrelation is indeed one. This example also shows that partial autocorrelation function can decrease rapidly given the right constellation of parameter values.

The Equation 2.1 can be written in a much more concise form using the notation of Lag Operators. Specifically:

$$Y_t - \mu = (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)\varepsilon_t$$

Figure 2.2: Example of a Moving Average process



Source: author's computations.

Definition 2.6 (Moving Average Operator). We define Moving Average Operator $\theta(L)$ as

$$\theta(L) = (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)$$

Quite important for MA processes is also the problem of uniqueness of the series. To give an example, MA(1) process with $\theta = 5$ and $\sigma^2 = 1$ would be identical to a process with parameters $\theta = 1/5$ and $\sigma^2 = 25$, provided that the underlying realization of the White Noise process would be identical. This "identification problem" is important for estimation and we also have to keep this issue in mind in further analysis.

2.2.3 Autoregressive Process of Order p

A natural extension of our discussion would be to allow the current value of $\{y_t\}$ to depend directly on its own past values and not just on the values of the White Noise. This is accomplished by the Autoregressive Process.

Definition 2.7 (Autoregressive Process of Order p). We define Autoregressive Process of Order p, AR(p), in the form:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t, \tag{2.2}$$

in which the $\{\varepsilon_t\}$ is a White Noise process from Definition 2.4. $\phi_1, \ldots, \phi_p \in \Re, \phi_p \neq 0, p \in \mathbb{N}$.

To simplify the notation:

Definition 2.8 (Autoregressive Operator). We define Autoregressive Operator $\phi(L)$ as

$$\phi(L) = (1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)$$

AR(p) can thus be stated in concise form as:

$$\phi(L) Y_t = c + \varepsilon_t \tag{2.3}$$

Now, if we replace the Lag Operators in the Autoregressive Operator by some variable, say x, the resulting polynomial will indicate explosiveness of the model. Specifically, AR(p) will be stable, if all of the roots of this polynomial will lie outside the unit circle, that is, if they will be greater in absolute value than one. We are not interested in explosive models for several reasons, one of them is that in theory, if the explosive process started at $t = -\infty$, it would not have any sensible values at the time of observed sample. More importantly, a model which would predict infinite growth would in a typical situation not be the most likely one.

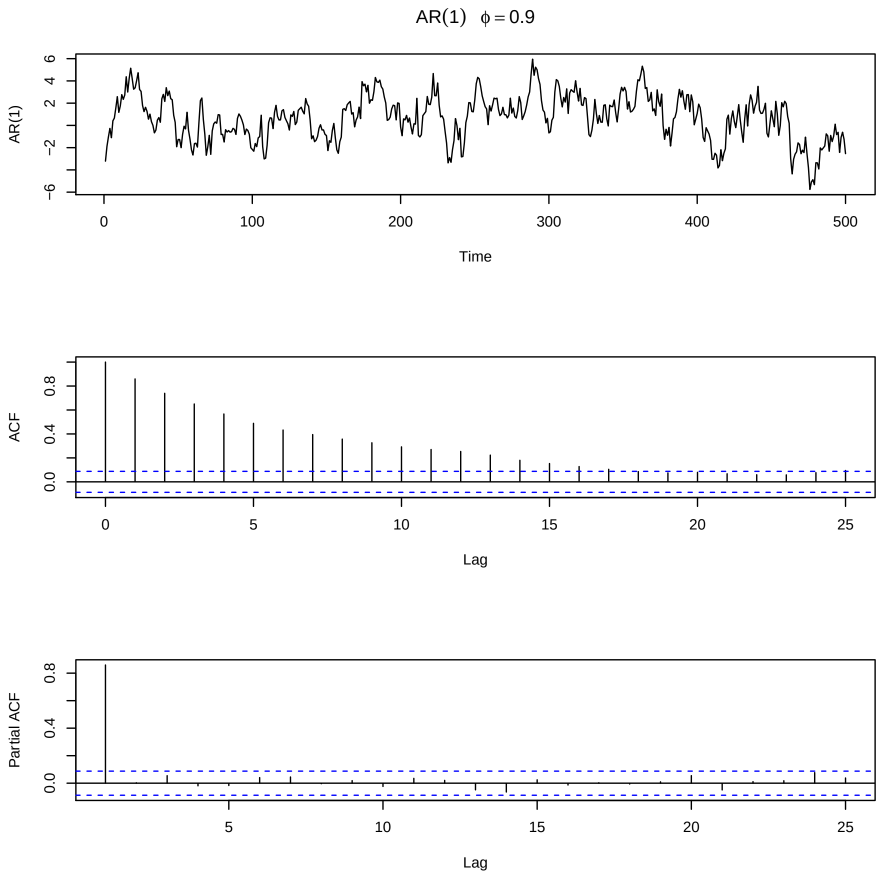
It can be shown that unconditional mean of the process of (stationary) AR(p) process is $\mu = \frac{c}{1-\phi_1-\ldots-\phi_p}$.

Theoretical autocorrelation function is computed in a similar way as that for Moving Average. The result for AR(1) is:

$$\rho_h = \phi^h$$

This process therefore does not have long memory from the 2.3 because the series $\sum_{h=0}^{\infty} |\phi|^h$ converges if the root of the respective polynomial lies outside the unit circle (which is in this case equivalent to the condition of $|\phi| < 1$).

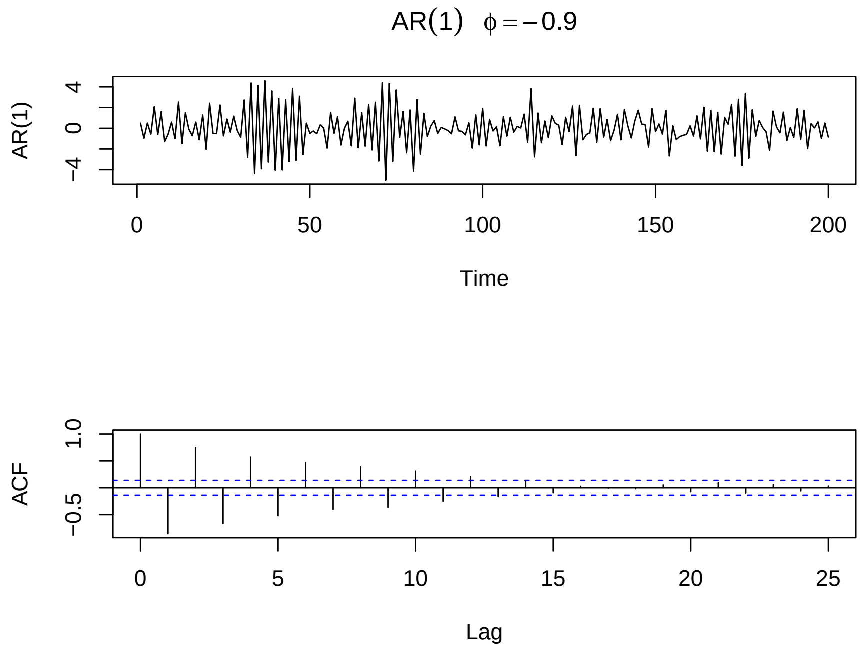
Figure 2.3: Autoregressive process of order 1



Source: author's computations.

There are two examples of AR(1) process included in this section. First

Figure 2.4: Mean Reverting Autoregressive process of order 1



Source: author's computations.

figure corresponds to a realization of a process $Y_t = 0.9Y_{t-1} + \varepsilon_t$ while the second figure is an example of mean reversion caused by negative impact of the last observation on the current one, $Y_t = -0.9Y_{t-1} + \varepsilon_t$. Several important observations can be made for the first figure. Firstly, as predicted, sample autocorrelation function does not have a clear cut-off point but decreases at an approximately exponential rate to zero. Secondly, there is precisely one significant lag in the sample partial autocorrelation function. It turns out that partial autocorrelation function of a stationary AR(p) model can have a non-zero values only at first p lags. Lastly, the value of the sample PACF at its first lag is equal to value of the ACF at its first lag.

There is an important connection between AR and MA processes. A stationary AR process is *causal* if it has an $MA(\infty)$ representation. To see this, consider Equation 2.3. This equation can be divided through by $\phi(L)$ and then, provided that the process is stationary, expanded with use of formula for sum of geometric series. In case of higher orders p, the fraction would first need to be separated into partial fractions in order to obtain the required form.

We call a MA process with $AR(\infty)$ representation *invertible*. The justifications and derivations are similar in principle.

2.2.4 ARMA processes

Definition 2.9 (ARMA(p,q)). The ARMA(p,q) is defined using previous definitions as:

$$\phi(L)(Y_t - \mu) = \theta(L)\varepsilon_t$$

This process is also stationary and without Long Memory if the roots of the polynomial associated with $\phi(L)$ lie outside the unit circle. Autocorrelations are slightly more complicated for the first q lags but afterwards they just return to standard AR(p) exponential decay.

Moreover, the polynomials $\phi(L)$ and $\theta(L)$ should not have common roots because that would be redundant. In estimation of this model ($AR(p,q)$) it can therefore sometimes happen that the estimated coefficients will yield roots that are close to each other. Estimating a $AR(p-1,q-1)$ model could be in this case a sensible simplification.

2.2.5 ARIMA(p,d,q) processes

Definition 2.10 (ARIMA(p,d,q)). The ARIMA(p,d,q) is defined using previous definitions as:

$$\phi(L)(1-L)^d(Y_t - \mu) = \theta(L)\varepsilon_t, \tag{2.4}$$

where the parameter $d \in \mathbb{N} + \{0\}$ for the moment.

The intuition behind this model is connected to differencing time series. In many applications, our time series of interest can suffer from trends or unit roots. Performing first difference on this time series often leads to a stationary time series which are more suitable for econometric inference. For example, first difference of a Random Walk process (defined for our purposes as $AR(1)$ with $\phi = 1$) is just White Noise. In much similar way, ARIMA(p,d,q) models generate time series that after d differences remain just simple ARMA(p,q).

2.2.6 ARFIMA(p,d,q) processes

Definition 2.11 (ARFIMA(p,d,q)). The ARFIMA(p,d,q) is defined using previous definitions as:

$$\phi(L)(1-L)^d(Y_t - \mu) = \theta(L)\varepsilon_t, \tag{2.5}$$

where the parameter $d \in (-0.5; 0.5)$.

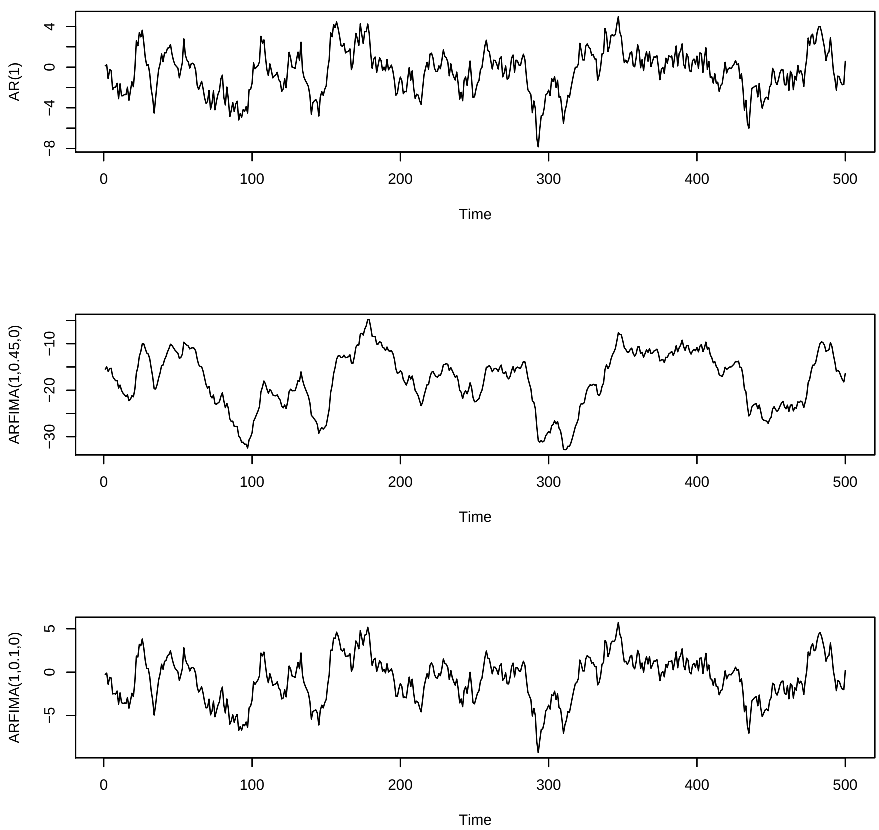
This extension of ARIMA models, originally introduced by Granger and

Joyeux (1980), Granger (1980, 1981), and Hosking (1981) (Baillie, 1996), is a standard tool in modelling Long Memory processes. This model exhibits Long Memory from Definition 2.3 for $d \in (-0.5; 0.5)$. If $d \in (-0.5; 0)$ however, the term *anti-persistence* is used as low values (not in absolute terms) are followed by large and vice versa.

Moreover, ARFIMA(p,d,q) is a truly general model as it successfully incorporates all of the previously mentioned model.

The following figure provides an example of three models: AR(1), ARFIMA(1,0.45,0) and ARFIMA(1,0.1,0). The autoregressive parameter was equal to 0.9 in all three cases and we used the same underlying White Noise sample. High persistence of the series on the second graph is visible even to a "naked eye" while the series in third graph can be hardly differentiated from the first one without proper statistical tools.

Figure 2.5: Example of ARFIMA processes



Source: author's computations.

2.3 Processes in the GARCH form

2.3.1 ARCH model

Definition 2.12 (ARCH(q) model)). We say that y_t follows ARCH (Autoregressive Conditional Heteroscedasticity) model of order q if:

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha_1 Y_{t-1}^2 + \cdots + \alpha_q Y_{t-q}^2 \\ Y_t &= \sigma_t \varepsilon_t, \end{aligned} \tag{2.6}$$

where $\varepsilon_t \sim N(0,1)$ and $\alpha_0 > 0, \alpha_i \geqslant 0$, for $i = 1, 2, \ldots, q$.

This model was introduced by Engle (1982) and has played an important role in development of volatility modelling in finance. The basic novelty of ARCH models was to allow variance (as a measure of risk) to "result from a specific type of non-linear dependence rather than exogenous structural changes in variables" (Bera & Higgins, 1993, p. 315). This meant that changes in variance of the time series over time could be properly modelled and taken care of within the time series itself.

Using simple algebra, the ARCH(1) model can be rewritten as:

$$Y_t^2 = \alpha_0 + \alpha_1 Y_{t-1}^2 + v_t, \tag{2.7}$$

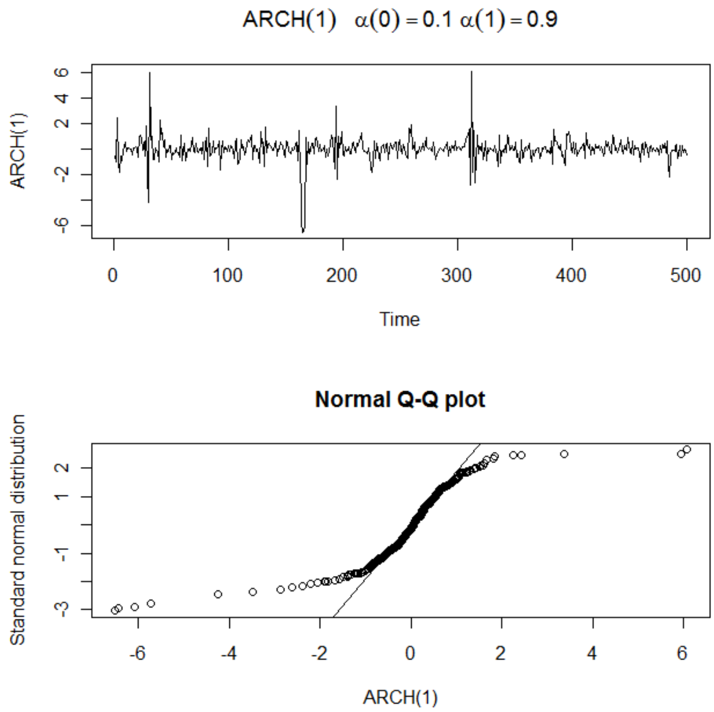
where usual conditions apply and $v_t = \sigma_t^2(\varepsilon_t^2 - 1)$. $v_t \sim \chi_1^2$ and is shifted by 1 to have zero mean. More importantly, we can see from the initial 2.12 that conditional distribution of Y_t on Y_{t-1} follows normal distribution $Y_t|Y_{t-1} \sim N(0, \alpha_0 + \alpha_1 Y_{t-1}^2)$ with volatility dependent on the past values of the time series. Unconditional variance turns out to be $Var(Y_t) = \mathbb{E}(Y_t^2) = \frac{\alpha_0}{1-\alpha_1}$, with an additional assumption that $\alpha_1 < 1$.

ARCH models gained wide appeal also because they are able to exhibit another typical property of financial time series, namely, fat tailed distribution of returns. One possible measure of fat tails is kurtosis, for ARCH(1):

$$K = \frac{\mathbb{E}(Y_t^4)}{\mathbb{E}[(Y_t^2)]^2} = 3\frac{1-\alpha_1}{1-3\alpha_1^2}$$

For $\alpha_1 < \sqrt{1/3}$ is $K > 3$ and therefore the resulting distribution is, by definition, leptokurtic with fat tails and high peak at mean value. For $\alpha_1 \geqslant \sqrt{1/3}$ however, the estimated series also exhibit fat tails and perhaps even more persuasive, as seen on Q-Q plot against standard normal distribution. This apparent disparity may be caused by inaccuracy in using kurtosis as a measure for fat tails. One of the reasons why this measure may be inadequate is that "it cannot account for peakedness and fat tails separately" (Schmid & Trede, 2003, p. 1).

Figure 2.6: ARCH(1) model



Source: author's computations.

Figure 2.6 provides a sample realization of ARCH(1) model,$\sigma_t^2 = 0.1 + 0.9Y_{t-1}^2$ with $Y_t = \sigma_t \varepsilon_t$. We can see how this model indeed generates fat tails so often present in financial time series. This stylized fact has important practical implications. It means that extreme gains or losses are far more likely than the classical the assumption of normally distributed returns would imply which in turn affects traditional models of volatility in risk management or various Value at Risk models by affecting the probability distribution of possible losses.

2.3.2 GARCH model

Definition 2.13 (GARCH(p,q) model)). We say that y_t follows GARCH (Generalized Autoregressive Conditional Heteroscedasticity) model of orders p, q if:

$$\sigma_t^2 = \alpha_0 + \alpha_1 Y_{t-1}^2 + \cdots + \alpha_q Y_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2$$

$$Y_t = \sigma_t \varepsilon_t,$$

where $\varepsilon_t \sim N(0, 1)$ and other reasonable conditions.

This extension of ARCH models was developed by Bollerslev (1986) and proved remarkably robust in practical applications. Even simple GARCH(1,1) model performed well for diverse financial time series, as pointed out by Engle in his Nobel Prize lecture. "It is remarkable that (GARCH(1,1)) can be used to describe the volatility dynamics of almost any financial return series" (Engle, 2003, p. 5). For this reason, simulated time series in this paper use GARCH model for innovations in order to better approximate real life time series.

The breadth of distinct models based on GARCH framework is staggering. Engle (2003) offers a handful, such as FIGARCH, designed specifically for fractionally integrated processes, EGARCH, used to account for different responses to positive and negative shocks, TGARCH, operating with similar idea, and a great deal of others.

Chapter 3

Statistical Tests for Long Memory

The idea of long memory in time series came from empirical work. This means that first efforts to detect long memory were heuristic in nature, as the proper definition of the term long memory had not been put down yet which means that they stemmed from experience as opposed to axiomatic deduction. These techniques can be therefore considered to be merely "simple diagnostic tools" (Beran, 1994, p. 81). But despite their drawbacks, they can provide us with first hints or general idea about whether we need to deal with long memory with some more sophisticated tools.

3.1 Heuristic estimation

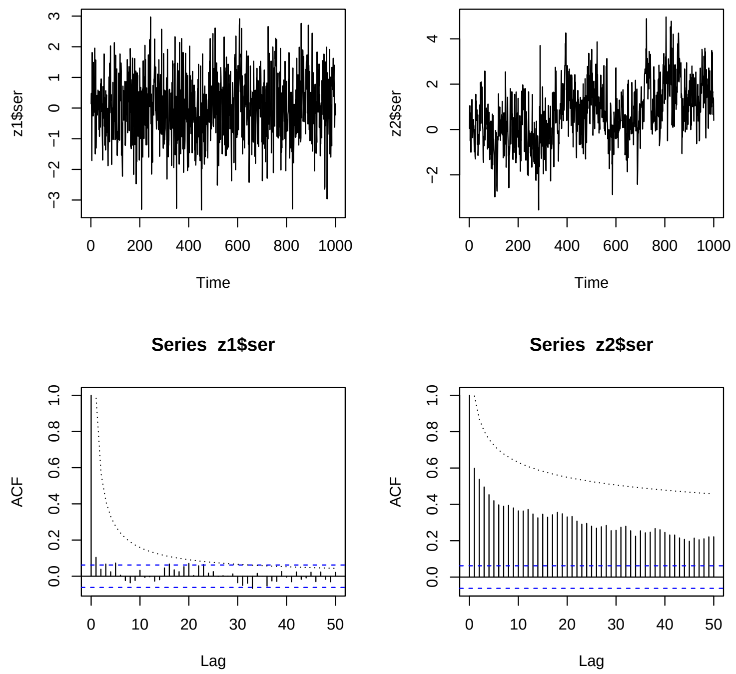
3.1.1 Autocorrelation function (ACF) and partial autocorrelation function (PACF)

We have used this simplest of the methods implicitly throughout the text because of its intuitive similarity with our definition of long memory, Definition 2.3.

It turns out that the number of significant lags in ACF is a good basis for setting the order of Moving Average process and, analogously, that number of significant lags in PACF can help us determine the order of the underlying Autoregressive process. But despite the similarity with our definition, the following example shows why we generally do not want to base the conclusion of whether the series has or has not long memory from the intuitive interpretation of ACF or PACF.

Figure 3.1 provides true and simulated ACF or ARFIMA(0,0.1,0) (on the

Figure 3.1: Difference between theoretical (dashed) and estimated (bars) ACFs of two ARFIMA processes.



left) and ARFIMA(0,0.4,0) (on the right) processes. Both of the series estimated in the figure have long memory, but the problem is that in case of ARFIMA(0,0.1,0) the estimated and theoretical autocorrelations are close to zero. This obviously does not affect the long memory, as even the series with very low absolute values can be divergent. Moreover, 95% confidence bands, as I have noted in Chapter 2, are valid for White Noise but not in general for any underlying stochastic process. This means that we can not with a sufficient degree of certainty rely on ACF or PACF in checking the long memory of a time series.

3.1.2 Rescaled Range Statistic (R/S)

The Rescaled Range Statistic (R/S) by E. Hurst was the first tool developed to deal with long memory. It was applied to study of the Nile River level data which had been long known for long periods of high and long periods of low water level without cyclical behaviour that would be clearly apparent. The construction of the R/S statistic works as follows:

1. Calculate cumulative sum series: $Z_t = \sum_{i=1}^{t}(Y_t - \bar{Y})$, for $t = 1, \ldots, n$

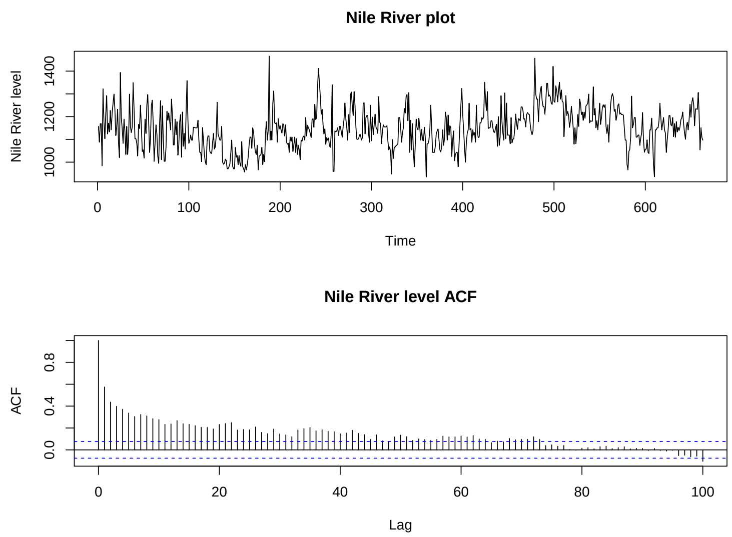
2. Calculate range series: $R_t = max(Z_1, Z_2, \ldots, Z_t) - min(Z_1, Z_2, \ldots, Z_t)$, for $t = 1, \ldots, n$

3. Calculate standardization series: $S_t = \sqrt{\frac{1}{t}\sum_{i=1}^{t}(Y_i - \bar{Y}_t)^2}$, for $t = 1, \ldots, n$ and where $\bar{Y}_t$ is average of observations Y_1 through Y_t.

4. The actual statistic is just R divided by S.

The R_t range is the ideal capacity of a reservoir which has uniform outflow, the water level is independent of t and the reservoir never overflows.

Figure 3.2: Yearly minimum water levels of the Nile River during 622-1284 measured at the island of Roda, near Cairo, Egypt



Source: http://mldata.org/repository/data/viewslug/nile-water-level/.

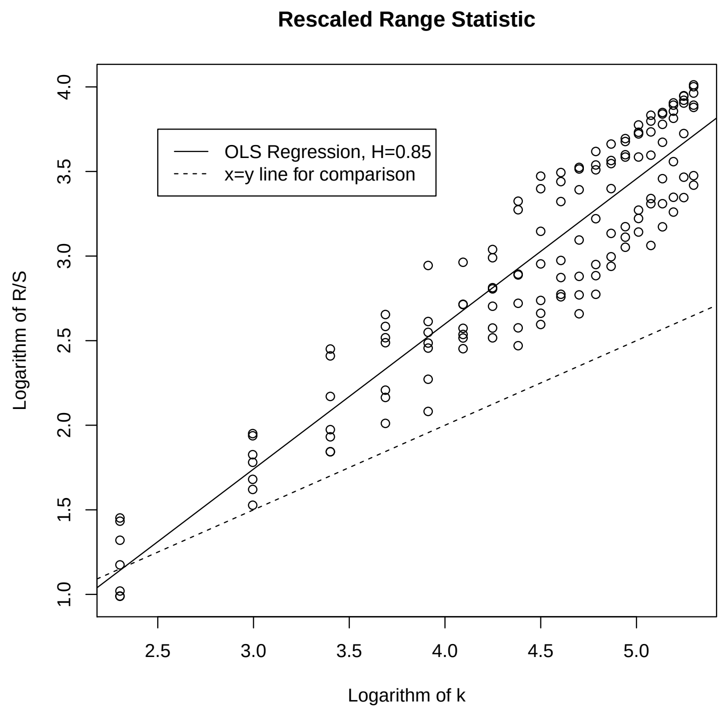
It is possible to calculate this statistics for different starting values of t as well as ending n. To arrive at an actual estimate of the long memory, one needs to estimate a regression of logarithm of the R/S statistics on the logarithm of the respective n used in the estimation of the statistics: $log(R/S) = \alpha + \beta \cdot n + \varepsilon$. The coefficient β is called Hurst parameter (H) and, in theory, $H \in (0, 1)$. It models rate of decay of the autocorrelation function proportional to k^{2H-2}. This implies that values greater than 0.5 indicate presence of long memory and values lower than 0.5 would indicate antipersistence. Unfortunately, the drawback of this test is that it does not sufficiently well distinguishes between long memory and other forms of dependences. Specifically, the test might show high values of the parameter H in case of slowly decaying time trends (Beran, 1994, p. 85).

Moreover, the theoretical results regarding distribution of the H statistic are not completely satisfactory, which makes it difficult to perform statistical inference. In particular, the distribution of H seems to depend on the underlying data generating process, choose of t and n, and even the length of the sample, see (Murphy & Izzeldin, 2000), (Barunik & Kristoufek, 2010), (Kristoufek, 2012) and (Weron, 2002).

It is important to note an important connection between Hurst parameter and parameter d in ARFIMA models. As was proved by Geweke & Porter-Hudak (1983) for ARFIMA models, $d = H - 1/2$.

Figure 3.3 is an example of estimation of the H parameter in the Nile River data (with $t = 60m + 1$, for $m = 1 \dots 7$ and $n = 10l$, for $l = 1 \dots 20$, following Beran (1994)). As expected, the estimate is well above the 0.5 level which is an indication of long memory present in the data.

Figure 3.3: Estimation of H parameter of the yearly minimum water levels of the Nile River



3.1.3 Modified Rescaled Range Statistic (R/S)

The modification of the Rescaled Range Statistic by Lo (1991) was developed in order to remedy the shortcomings of the original R/S statistic, especially its lack of well-defined distribution. In short, the difference lies in use of "consistent estimator of the long-run standard deviation, such as Newey-West(1987) estimator", (Murphy & Izzeldin, 2000, p. 352). The estimate of standard deviation takes into account the covariances of the first q lags (Teverovsky *et al.*, 1999). Under other mild conditions, the distribution of the Modified Rescaled Range Statistic asymptotically converges to a well-defined distribution, however, according to (Murphy & Izzeldin, 2000, p.), the finite sample distribution of R/S is not well approximated by its asymptotic distribution even when T is large.

Moreover, Kristoufek (2012) reports significant downward bias in Modified Rescaled Range estimation of parameter H in a Monte Carlo study. This would imply that Modified R/S test is biased towards rejecting long range dependence.

3.1.4 Detrended Fluctuation Analysis (DFA)

This improvement of classical Fluctuation Analysis (FA) was introduced by Peng *et al.* (1994) and, according to Grau-Carles (2006), is supposed to deal with power-law correlations in non-stationary time series. This is accomplished by performing linear or higher polynomial order time detrending of the time series in several non-overlapping intervals separately. The order of polynomial used is sometimes denoted in terms of $DFA1$ for linear trend, $DFA2$ for quadratic and so on. Unfortunately, no asymptotic distribution of this statistic has been discovered so far (Grau-Carles, 2006). The construction of this test is similar to that of R/S test:

1. Calculate cumulative sum series: $Z_t = \sum_{i=1}^{t}(Y_t - \bar{Y})$, for $t = 1, \ldots, n$
2. Divide the whole set into k non-overlapping intervals with m observations in each and perform least squares regression of Z_t on a (linear or higher polynomial order) function of time.
3. Calculate the fitted values from these regressions $\hat{y}_{mt}$.
4. Compute $F_m = \sqrt{\frac{1}{m \cdot k} \cdot \sum_{i=1}^{m \cdot k}[y_t - \hat{y}_{mt}]^2}$ for several values of m and k.
5. Regress $log(F_m)$ on $log(m)$ and estimate the slope parameter γ by OLS.

The slope parameter γ has similar interpretations as the Hurst parameter H. γ equal to 0.5 indicates no long memory, values higher than 0.5 indicate long memory and if γ is lower than 0.5 then we may face long term anticorrelation or antipersistence. What is different however, due to detrending, we can also interpret values higher than 1 as an indication of non-stationarity of the data cause by deterministic or stochastic trends (which we differenced-away in the estimation process).

In Figure 3.4 and Figure 3.5 we perform a sample DFA test for clarity reasons. With $\hat{\gamma} = 0.95$, DFA test also supports the conclusion that Nile River level data exhibit long range dependence.

Figure 3.4: DFA performed on the Nile River level data

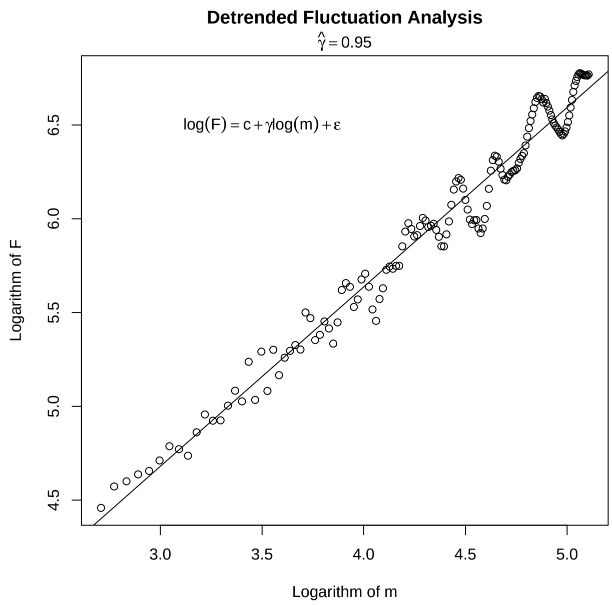
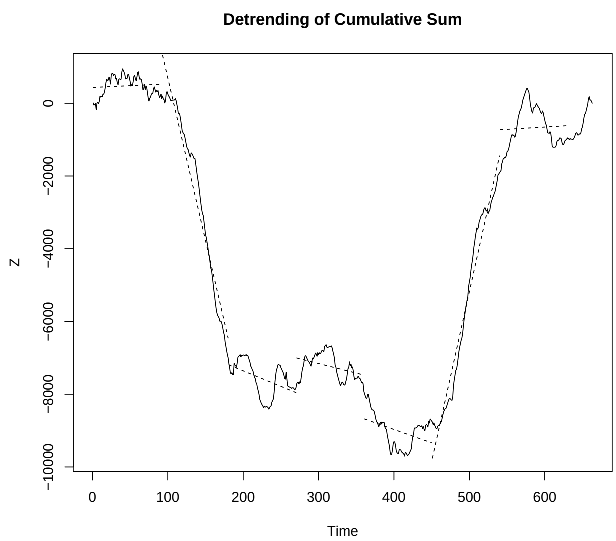


Figure 3.5: Detrending with $m = 100$



3.2 Time and Frequency Domain Estimation

3.2.1 Exact Maximum Likelihood Estimation (MLE)

Heuristic methods discussed in the previous sections are useful especially when we are interested in finding out whether $H > 1/2$ or nor, that is, whether we have a time series with or without long memory. They are, however, not well suited for statistical analysis and estimation of the variance of these estimators, yet alone their distribution, is not easy. Under some additional assumption we are able to build Maximum Likelihood Estimators that can to a certain degree counteract these issues. Most importantly, Beran (1994) stresses the fact that MLE is clearly more efficient than the previous methods, provided that we can build a reasonable parametric model. A parametric model is simply an assumption about the form of the underlying process and from that the joint density function of our observations in the time series. We will restrict our discussion to Gaussian Likelihood, primarily because of its simplicity, as normal distribution is fully specified by its first two moments only. This does not imply that our results will be valid for Gaussian time series exclusively, in other words, the results will hold in more general cases as well.

We will want our time series vector $Y = Y_t$, for $t = 1, \ldots, T$, to be a realization of causal invertible process, which means that it has an $MA(\infty)$ and $AR(\infty)$ representation. This implies that Y_t depends on its past values and that the dependence is linear. While the former implication seems intuitively justifiable, the latter is potentially serious simplification assumed for computational purposes. We define $\Sigma_T(\theta)$ to be a covariance matrix of Y, $|\Sigma_T|$ is determinant of the matrix, θ is the parameter vector to be estimated. We can simplify our computation by setting mean of the process to be just simple average, that is $\mu = 1/T \cdot \sum_{i=1}^{T} Y_i$. We can write the joint density as:

$$h(y; \theta) = (2\pi)^{-n/2} \cdot |\Sigma_T(\theta)|^{-1/2} \cdot exp\{-\frac{1}{2} Y^T \cdot \Sigma_T^{-1}(\theta) \cdot Y\}$$

And log likelihood as:

$$\mathcal{L}_T(y; \theta) = log[h(y; \theta)] = -\frac{n}{2} log(2\pi) - \frac{1}{2} log|\Sigma_T(\theta)| - \frac{1}{2} Y^T \cdot \Sigma_T^{-1}(\theta) \cdot Y$$

We would now want to find the maximum of $\mathcal{L}_n(y; \theta)$ with respect to θ. This is done by setting $\nabla \mathcal{L}_n(y; \theta) = 0$. The resulting system of equations is however complicated, so one usually resorts to approximate methods, such as Whittle

Estimation.

3.2.2 Whittle's Approximate Maximum Likelihood Estimation (Whittle's MLE)

One of the possible simplifications of the Exact MLE is the Whittle MLE. This method lies in, first, approximating the Exact MLE likelihood function with the following expression:

$$\mathcal{L}_W(\theta) = \frac{1}{2\pi} \cdot \int_{-\pi}^{pi} log f(\lambda;\theta) d\lambda + \frac{Y^T A(\theta) Y}{T}$$

The expression is derived first by noticing that the first term in the Exact MLE does not depend on the parameter θ and than by approximation of the two other terms. Consult (Beran, 1994) or (Palma, 2007) and references herein for details. $f(\lambda;\theta)$ is spectral density of the process, $A(\theta)$ is $n \times n$ matrix with $a_{j,l} = (2\pi)^{-2} \int_{-\pi}^{\pi} \frac{1}{f(\lambda;\theta)} e^{i(j-l)\lambda} d\lambda$.

Discrete version of the estimator is derived by replacing integrals with Riemann sums and using periodogram of the process $I(\lambda)$:

$$\mathcal{L}_W(\theta) = -\frac{1}{2\pi} \left[\int_{-\pi}^{\pi} log f(\lambda;\theta) d\lambda + \int_{-\pi}^{\pi} \frac{I(\lambda)}{f(\lambda;\theta)} d\lambda \right]$$

$$\mathcal{L}_D(\theta) = -\frac{1}{2T} \left[\sum_{j=1}^{T} log f(\lambda_j;\theta) + \sum_{j=1}^{T} \frac{I(\lambda_j)}{f(\lambda_j;\theta)} \right]$$

3.2.3 Geweke and Porter-Hudak Estimator (GPH Estimator)

GPH Estimator was introduced by Geweke & Porter-Hudak (1983) and they also derived its asymptotic distribution. The advantage of this approach is its relative simplicity, the estimation procedure in performed by least squares regression. The estimation of the parameter d from ARFIMA(p,d,q) model is consistent using the *GPH* Estimator (Murphy & Izzeldin, 2009). The regression equation comes down to:

$$log(I(\lambda_j)) = c - d \cdot ln(4 sin^2(\lambda_j/2)) + \varepsilon_j, \quad j = 1, \ldots m,$$

where $I(\lambda_j)$ is the periodogram of the Y_t time series at frequencies $\lambda = 2\pi/T$. m is often set equal to $\lfloor T \rfloor$. This estimator is asymptotically unbiased and has variance equal to $\pi^2/6$, $\hat{d} \sim N(d, \pi^2/6)$ (Murphy & Izzeldin, 2009).

3.2.4 A Wavelet based approach

A proper treatment of wavelets would require a lengthy detour, as the topic is broad with many diverse applications in several fields ranging from image processing to particle physics. In general however, wavelets can be used to derive certain properties from the data, such as present of long memory. A wavelet is a function ψ that satisfies the following conditions:

(1) $\int_{-\infty}^{\infty} \psi(u)du = 0$

(2) $\int_{-\infty}^{\infty} \psi(u)^2 du = 1$

Wavelets come in a variety of forms and shapes, and are generally differentiated into two groups (or waves): the first wave resulted in continuous wavelet transformation, which deals with time series defined over the entire real axis, and the discrete wavelet transform (DWT) (Percival & Walden, 2000). In the study of empirical time series, one is often required to focus on the techniques from the discrete wavelet transform group.

The estimation of long memory is statistically problematic due to a high degree of correlation among the variables. DWT creates new random variables, denoted d_{jk}, that are approximately uncorrelated. That is done by:

$$d_{jk} = \int_{-\infty}^{\infty} y(t)\psi_{jk}(t)dt,$$

where $\psi_{jk}(t)$ is in the form $2^{-j/2}\psi(2^{-j}t - k)$ and j is usually called an octave. Under some other regulatory conditions we have decomposed the original process y_t into double sum:

$$y(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_{jk}\psi_{jk}(t)$$

There are several suitable candidates for the ψ function, such as Haar or Daubechies wavelets (Palma, 2007).

The wavelet based approach to estimating long memory was introduced by Jensen (2000). To estimate the long memory parameter d, one needs to first calculate the mean of estimated d_{jk} for each j, $\hat{\mu_j} = \frac{1}{n_j}\sum_{k=1}^{n_j} \hat{d}_{jk}^2$. Then regress this variable on scale parameter with heteroscedasticity robust standard errors. For the estimated coefficient at the scale, $\hat{\beta}$, holds that $\hat{\beta} = 2\hat{d}$.

3.2.5 Other methods

Time series literature is abundant with tests for long memory. To mention at least some of them, KPSS can be used in the long memory estimation as well, Robinson's $\hat{H}$, $\hat{S}_k$, and there is also a whole separate theory of wavelets, see (Murphy & Izzeldin, 2009) for estimators and (Percival & Walden, 2000) for an extensive treatment of wavelets.

3.3 Moving Block Bootstrap (MBB)

The idea of using bootstrap in time series analysis comes from the fact that this procedure can give us an estimated small sample distribution of the statistic we are interested in. It uses computationally intensive method of resampling from the original time series to replicate the sample which allows us to obtain a whole set of estimates. We can use this sample to construct histograms or p-value of the estimate from the original distribution.

Although the original bootstrap method was developed by Efron (1982), this technique fails when the observations are not independent, identically distributed (i.i.d) (Singh, 1981). Kuensch (1989) developed a modification of the bootstrap apt for non i.i.d. observations. While in Künsch's version, resampling to create the bootstrapped sample is performed from separate sets with replacement, in the Moving Block Bootstrap (MBB) by Davidson & Hinkley (1997) we allow for overlapping sets. This creates greater variety of our bootstrapped sample especially when the block length is long.

There are several theoretical problems associated with this technique. Murphy & Izzeldin (2009) mention that bootstrapped samples may not be stationary even if the original process was. Despite the concerns, the MBB provides reasonable results in practise.

We implement the "post-blacken" MBB for our estimated time series y_t with the following structure:

Step 1. "Pre-whiten" our time series y_t by fitting an $AR(p)$ model and obtain the residuals e_t.

Step 2. Estimate centered residuals $r_t = e_t - \bar{e}$

Step 3. Resample n blocks from r_t of length k, so that, preferably, $n \cdot k = T$.

Step 4. "Post-blacken" the resampled series by fitting the original estimated AR(p) model with the original parameters.

Step 5. Compute the statistic of interest and repeat the relevant steps until you receive a sufficiently smooth distribution of the estimates.

We use Schwarz criterion (SC) to obtain the lag order p in the $AR(p)$ model but there are also other criteria which might also be considered, especially Akaike (AIC) and Hannan-Quinn (HQ). Both SC and HQ are consistent estimators of the true lag order, AIC is actually inconsistent as it overestimates the true lag order. On the other hand AIC may be more efficient in finite samples. For $T \geq 16$ holds that $AIC \geq HQ \geq SC$, where, as a shorthand, the name represents number of lags suggested by the respective criterion. See Brueggemann & Leutkepohl (2000) for further details.

Estimation of the model can be quite different than its generation. There are several methods for estimation of $AR(p)$ parameters, notably Yule-Walker, Least Squares estimation and three types of likelihood estimators: Conditional Maximum Likelihood, with fixed "pre-sample" values, Unconditional Maximum Likelihood, which obtains "pre-sample" values by "backcasting", and Exact Maximum Likelihood. The techniques yield similar results (Yule-Walker method is asymptotically equivalent to Least Squares) and we therefore use the Least Squares method because it is not demanding computationally.

Chapter 4

Monte Carlo Study

In the chapter we present the results of our Monte Carlo study and provide specifics of the procedures carried out.

4.1 Choice of Tests

We have performed the estimation for four long range dependence tests:

- R/S
- DFA
- GPH
- Wavelet based method

This choice reflects an effort to select most widely used tests in the literature. Notably, R/S and GPH estimators are often used in the applied work and so is DFA. Robinson & Henry (1999) mentions that GPH is one of the two leading semiparametric estimates of the memory parameter H. Wavelet based method is also a very promising method, given the amount of research in the theory of wavelets. This selection could naturally be expanded upon, especially when we consider the breadth of available methods.

When deciding whether to reject or not to reject the null of no long memory, we relied on "naïve" decision-based process by assuming normal distribution of the estimates of H. This approach is asymptotically justifiable in case of GPH, wavelet based method, and also R/S given no short memory but is problematic for DFA, as was explained in Subsection 3.1.4. It is nonetheless reasonable to perform this estimation to see just how relevant is the inaccuracy researchers would induce when relying on such an estimate.

In the R/S method, we divide the time series into 50 non-overlapping blocks. In the DFA method, we use linear detrending on the blocks of equal length as in the R/S method. For GPH, we use bandwidth parameter of the size equal to $\lfloor \sqrt{T} \rfloor$. We used time series with length expressible at power of 2 (512 and 1024) due to wavelet estimator which is forced to truncate the time series to the nearest power in case of exceeding number of observations. Order of our wavelet estimator was set to 2 while octaves were 2 and 8.

Although we have developed original procedures for estimating R/S and DFA methods when writing the theoretical parts of this paper, in the Monte Carlo study we relied on procedures provided in econometric packages "fArma" and "fracdiff" of R-project for all estimators. There are several reasons for this choice. Firstly, our results are more directly comparable with the use of standardized procedures. Secondly, the results are more reliable when using standard procedures. And lastly, the standardized procedures may utilize more efficient estimation procedures as far as efficient programming is concerned, with efficiency being one of the most significant limitations in Monte Carlo studies. In the bootstrapping estimation, we used "post-blacken" moving block bootstrap as described in Section 3.3 with 100 bootstrap replications.

Using 3 computers, a run time of 20 hours was needed to perform 200 Monte Carlo simulations. Long estimation time could be attributed to DFA estimator which seems to be the one most computationally demanding. Moreover, despite using fast Schwarz criterion estimation method, we still forced the algorithm to estimate several models exactly to arrive at a reliable results and mimic real-life estimation in which we do not know the exact lag order.

4.2 Choice of Processes

We have used the following 9 models to perform the study:

- White Noise
- GARCH(1,1) with $\alpha = 0.1$ and $\beta = 0.8$
- ARMA(1,1) with $\phi = 0.5$ and $\theta = 0.1$
- ARFIMA(0,0.25,0)

- ARFIMA(1,0.25,0) with $\phi = 0.25$
- ARFIMA(1,-0.25,0) with $\phi = -0.25$
- ARFIMA(0,0.25,0) with GARCH(1,1) innovations
- ARFIMA(1,0.25,0) with $\phi = 0.25$ with GARCH(1,1) innovations
- ARFIMA(1,-0.25,0) with $\phi = -0.25$ with GARCH(1,1) innovations

We have generated these pseudo-random time series of length 512 and 1024, with an initial "burn-in" period of 100 observations. This selection offers 3 fairly general classes of time series without long memory and 6 ARFIMA models with either long memory or anti-persistence. Such a setting allows us to infer about both the size and power properties of the estimators in a scenario which resembles applied work reasonably well. This is especially due to the use of GARCH(1,1) model which, as noted in Subsection 2.3.2, has found many applications in finance. Fractionally integrated processes, on the other hand, are also very often used tools to model long memory. We have included ARMA(1,1) model to make the estimators face a process with short memory as well.

4.3 Choice of Parameters

The choice of parameter values was also motivated by efforts to approximate real-life time series. In the case of GARCH(1,1), Baillie *et al.* (1996) note extreme degree of persistence of shocks to variance in empirical studies. This has led us to select values of α and β with their sum close to one because, as it turns out, that is the factor affecting the level of integration of GARCH processes (Harvey & Streibel, 1998). Similar considerations affected the choice of ARMA and ARFIMA parameter values.

4.4 Results

The complete results are summarized in Table A.1 through Table A.9 for the respective processes. The tables present empirical size and power properties as calculated from empirical rejection frequencies of the null hypothesis of no long range dependence. For each time series and each test statistic, we offer rejection frequencies based both on the approximation of respective asymptotic distributions and on the moving block bootstrap procedure.

Table A.1 through Table A.3 represent estimates of size properties of the tests when the underlying time series do not have long memory. We can see a consistently better performance of the MBB in comparison to asymptotic critical values for the R/S statistic in all 3 cases. The R/S estimator appears to be oversized, as it rejects the null hypothesis too often and MBB is able to decrease the rejection frequencies closer to the critical values. Its performance is however still consistently inferior to both GPH and Wavelet estimators for all time series on almost all critical levels. The GPH estimator has the best size properties from these tests, especially because it was able to provide solid results in the ARMA(1,1) case which was the main stumbling block of the Wavelet based estimator. DFA provided completely unreliable results by rejecting the null at times in close to a one hundred per cent of cases. Moreover, its performance was not improved by MBB.

Simulations of power properties of the estimators for ARFIMA processes with normal innovations, Table A.4 through Table A.6, show an inferior performance of the GPH estimator which does not reject the null as often as needed. In this case, bootstrapping can improve the test performance, which means that under the null of no long range dependence, asymptotic rejection frequencies of GPH are better while under the alternative, bootstrap improves the performance of GPH estimator, albeit not markedly. Other tests perform reasonably well but in the cases of R/S and DFA it can be attributed to their overall high tendency to reject the no long memory null. Interestingly, MBB fails in case of anti-persistence in Table A.6. Wavelet estimator does not offer an excellent size and power properties but is reasonably robust in both of these categories.

We can also see that size and power properties of the estimators are better for longer time series but the improvements are not excellent. Especially the GPH estimator in case of ARFIMA processes with normal innovations does not show convergence to asymptotically expected rejection frequencies. Surprisingly, for the time series without long memory, bootstrap seems to be decreasing in quality for longer time series. This could imply that the number of bootstrap replications needed to ensure proper estimates is dependent on the time series length.

Overall, GPH estimator performs better when innovations follow the GARCH(1,1) process but fails in the ARFIMA(1,-0.25,0) case. Bootstrap fails in the ARFIMA(1,-0.25,0) case as well but other estimators perform reasonably well. MBB did not live up to its expectations and provided little improvement of the rejection frequencies. This result can be to a certain extent attributed to a small number of bootstrap replications. For the R/S statistic, fitting of AR(p) model in the process of bootstrap estimation could affect the null hypothesis, thereby skewing the whole bootstrap distribution and eventually providing less reliable estimates. The Wavelet estimator has performed very well, especially due to its high power against ARFIMA processes with normal innovations.

Chapter 5

Long Range Dependence Analysis of SAX Index

5.1 Description of the Index SAX

Slovak Share Index is the main index of Bratislava Stock Exchange, the only officially licensed stock exchange in Slovakia. It is capital-weighted index with included dividend payouts and changes in the number of shares outstanding due to new issues of corresponding shares. As a comparison index, we chose a standard index S&P 500. The data we used run back to July 1995 for SAX and January 1950 in case of S&P 500 with around 4000 and 15000 observations respectively. Figure 5.1 plots levels of the two time series.

5.2 Estimation

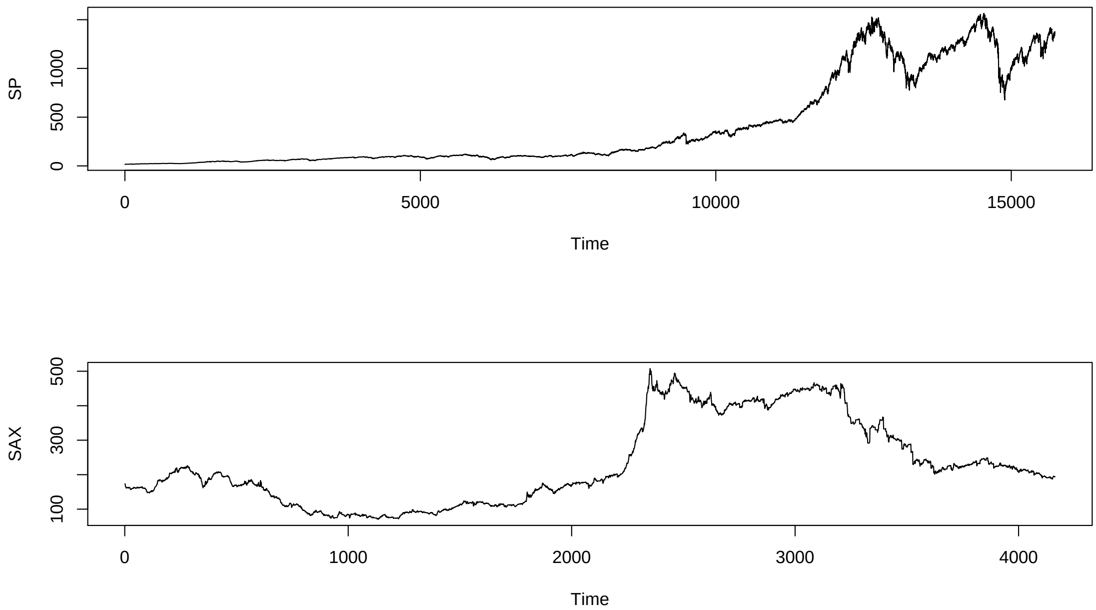
To receive growth rates, we use approximate formula

$$r_t = log(P_t) - log(P_{t-1}),$$

where P_t is level of index at time t, r_t is the growth rate or return at time t.

Figure 5.2 depicts growth rates of the two indices. The plots suggest presence of some form of autoregressive dependence and time-varying volatility in the two figures but we are interested in the presence of long memory. Several studies have supported the conclusion of presence of long memory in volatility of stock markets, (Henry, 2002) and (Kang & Yoon, 2007), while Kasman *et al.* (2009) provided mixed results for central and eastern Europe indices. In

Figure 5.1: Plot of SAX and S&P 500 indices



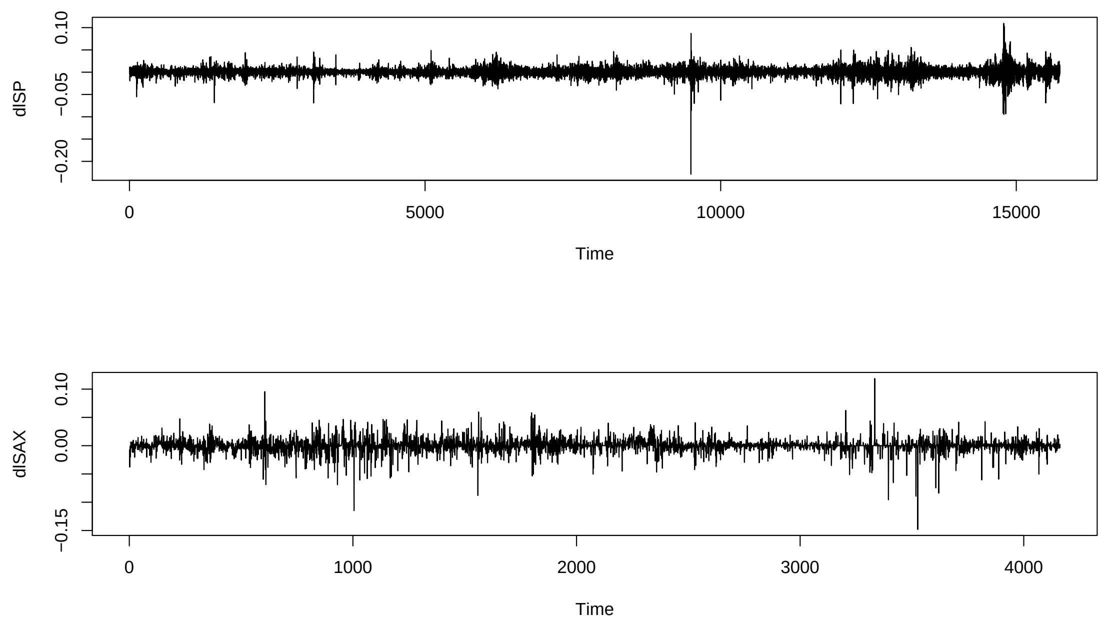
Source: `http://www.bsse.sk/obchodovanie/indexy/IndexSAX.aspx` for SAX and `finance.yahoo.com/q?s=^GSPC` for S&P 500.

particular, the study performed GPH test for SAX index and found evidence for LRD in levels and volatility.

We set out to test unit root in the logarithm of SAX level by Augmented Dickey Fuller test and KPSS. In case of ADF, both Schwarz criterion and Akaike criterion suggested lag order of zero and we have therefore performed simple Dickey Fuller test with constant term and time trend in the regression equation. With appropriate p-value of around 0.97%, we were not able to reject the null of unit root with drift. For differenced time series (e.g. returns) however, we can strongly reject unit root by Dickey Fuller test with constant term only (test statistic close to -65.92). In KPSS test we strongly rejected the null of trend stationarity but in the differenced time series, we were able to reject the null of simple stationarity on the ten percent level with shorter truncation. This makes the evidence not fully persuasive but we still conclude that the series is I(1) and we are free to perform the long memory tests on the first difference of the series.

We have arrived at crystal clear I(1) process in case of S&P 500 index when performing the same tests. Both differenced time series are plotted in Figure 5.2.

Figure 5.2: Plot of returns of SAX and S&P 500 indices



5.3 Results

The results of our study are presented in Table A.10. We used the same estimators and parameter settings as in our Monte Carlo study. Both GPH and R/S tests provide persuasive evidence for long memory in returns of SAX index which was anticipated given the results of Kasman *et al.* (2009). Bootstrap was able to provide a mild improvement in confidence levels of GPH estimator in our Monte Carlo study and its p-value of 0.01 therefore strengthens the plausibility of long memory hypothesis. On the other hand, Wavelet estimator did not reject the null of no long memory. It is a strong counter evidence given its performance in the Monte Carlo study but we consider the size of GPH test to be better than the power of Wavelet test and therefore tend towards rejecting the null.

We included S&P 500 index growth rates diagnostics for comparison and we can conclude that it does not exhibit long memory with high certainty. Moreover, magnitudes of estimated H are generally much closer to 0.5 than in the case of SAX index.

As far as volatility is concerned, S&P exhibits high and statistically significant long range dependence as seen in very small standard p-values for squared series, $S\&P^2$, and absolute value series, $|S\&P|$. This means that possible shocks to actual volatility may have long-lasting effects on the S&P 500 index. These results are in accordance with stylized facts about stock market in developed countries (Kirchler & Huber, 2009). Results for SAX index are so clear-cut. While the absolute value series seems to exhibit some persistence, p-values for squared returns are higher especially for GPH and Wavelet. This leaves question of long memory in volatility of SAX index open.

The bootstrap p-values turned out to be mostly unreliable in this study. This should however not come as a surprise considering that very high length of the data might have favoured asymptotic argumentation especially for the GPH and Wavelet estimators.

Chapter 6

Conclusion

We have devoted a significant part of this thesis to the development of the time series theory needed to study long range dependence in the AFRIMA framework. We then moved on to statistical tests for long range dependence and provided their careful descriptions and possible drawbacks. Those include but are not limited to either unknown (DFA) or only asymptotic distributions of the estimators with arguably poor finite sample approximations. One of the proposed remedies to this situation is the moving block bootstrap, a modification of original bootstrap by Efron (1982) for time series. The idea of the moving block bootstrap is to resample blocks of equal length from the original time series with replacement and estimate one value of a statistic of interest for each of these new time series. This randomized procedure then provides us with a whole set of estimates which can than be used for statistical hypotheses testing. Theoretically, the null of moving block bootstrap for any statistic is no long range dependence due to the short length of a block relative to the overall length of the time series.

This bachelor thesis aimed to provide Monte Carlo comparison of asymptotic and bootstrapped size and power properties of long range dependence tests. We evaluated performance of 4 estimators, R/S, DFA, GPH, and Wavelet based method, against 3 classes of time series with altogether 9 models. This was a fairly general setting that exposed the tests to quite a high variety of time series. The study revealed that moving block bootstrap can improve size of R/S test but that it does not offer reliable results in general. Especially in cases of GPH and Wavelet estimators, asymptotic standard errors provided more reasonable confidence intervals. The GPH estimator had better properties for processes without long memory than Wavelet but the Wavelet estimator was more robust to some long memory processes. The study also exposed DFA as a very unreliable estimator, even with moving block bootstrap. It is thus not suited for statistical inference and hypotheses testing.

We have applied the methods to the assessment of SAX index and S&P 500 index. Our study supported the idea of no long range dependence in the returns of S&P 500 and provided persuasive evidence for the presence of long memory in the volatility of the index. In case of SAX, we were able to replicate the results of Kasman *et al.* (2009) as the GPH test provided evidence for long memory in returns. Despite this result, the Wavelet estimator did not reject the null of no long memory on reasonable levels and considering good performance of this estimator in the Monte Carlo, we may question the validity of the previous conclusion. A resolution of this contrasting evidence is important because absence of long memory can be linked to efficiency of the market.

The most natural extension of this study would be to include different long range dependence tests, to perform the tests on different time series models or to adjust the values of parameters in the models and continue in the Monte Carlo fashion. We have postulated that the weak performance of moving block bootstrap might be due to the low number of bootstrap replications and it would be interesting to see whether it holds and what is the minimum number of replications needed to achieve a certain level of accuracy.

Bibliography

Baillie, R. T. (1996): "Long memory processes and fractional integration in econometrics." *Journal of Econometrics* **73(1)**: pp. 5–59.

Baillie, R. T., T. Bollerslev, & H. O. Mikkelsen (1996): "Fractionally integrated generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* **74(1)**: pp. 3 – 30.

Barunik, J. & L. Kristoufek (2010): "On hurst exponent estimation under heavy-tailed distributions." *Physica A* **389**: pp. 3844–3855.

Bera, A. K. & M. L. Higgins (1993): "Arch models: Properties, estimation and testing." *Journal of Economic Surveys* **7**: pp. 307–366.

Beran, J. (1994): *Statistics for long-memory processes.* Monographs on statistics and applied probability, 61. Chapman & Hall.

Bollerslev, T. (1986): "Generalized autorregressive conditional heteroskedasticity." *Journal of Econometrics* **31**: pp. 307–327.

Brueggemann, R. & H. Leutkepohl (2000): "Lag selection in subset var models with an application to a u.s. monetary system." *Econometric Society World Congress 2000 Contributed Papers 0821*, Econometric Society.

Davidson, A. & D. Hinkley (1997): *Bootstrap Methods and Their Application.* Cambridge University Press.

Ding, Z., C. W. Granger, & R. F. Engle (1993): "A long memory property of stock market returns and a new model." *Journal of Empirical Finance* **1(1)**: pp. 83 – 106.

Efron, B. (1982): "The jackknife, the bootstrap, and other resampling plans."

ENGLE, R. F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation." *Econometrica* **50**: pp. 987–1007.

ENGLE, R. F. (2003): "Prize lecture." Nobelprize.org as of 10 June 2012 `http://www.nobelprize.org/nobel_prizes/economics/laureates/2003/engle-lecture.html`.

GEWEKE, J. & S. PORTER-HUDAK (1983): "The estimation and application of long memory time series models." *Journal of Time Series Analysis* **4(4)**: pp. 221–238.

GRAU-CARLES, P. (2006): "Bootstrap testing for detrended fluctuation analysis." *Physica A* **360**: pp. 89–98.

HAMILTON, J. D. (1994): *Time-series analysis*. Princeton Univerity Press, 1 edition.

HARVEY, A. & M. STREIBEL (1998): "Testing for a slowly changing level with special reference to stochastic volatility." *Journal of Econometrics* **87(1)**: pp. 167–189.

HENRY, M. & P. ZAFFARONI (2002): "The long range dependence paradigm for macroeconomics and finance." *Discussion Papers 0102-19*, Columbia University, Department of Economics.

HENRY, O. T. (2002): "Long memory in stock returns: some international evidence." *Applied Financial Economics* **12(10)**: pp. 725–729.

HERNANDEZ-CAMPOS, F., L. LE, J. S. MARRON, J. PARK, V. PIPIRAS, F. D. SMITH, R. L. SMITH, M. TROVERO, & Z. ZHU (2011): "Long range dependence analysis of internet traffic." *Journal of Applied Statistics* **38**: pp. 1407–1433.

HURST, H. E. (1951): "Long-term storage capacity of reservoirs." *Trans. Amer. Soc. Civil Eng.* **116**: pp. 770–808.

JENSEN, M. J. (2000): "An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets." *Journal of Economic Dynamics and Control* **24(3)**: pp. 361 – 387.

Kang, S. H. & S.-M. Yoon (2007): "Long memory properties in return and volatility: Evidence from the korean stock market." *Physica A: Statistical Mechanics and its Applications* **385(2)**: pp. 591 – 600.

Kantelhardt, J. W. (2009): *Encyclopedia of Complexity and Systems Science: Fractal and Multifractal Time Series*. Springer.

Kasman, A., S. Kasman, & E. Torun (2009): "Dual long memory property in returns and volatility: Evidence from the cee countries' stock markets." *Emerging Markets Review* **10(2)**: pp. 122 – 139.

Kirchler, M. & J. Huber (2009): "An exploration of commonly observed stylized facts with data from experimental asset markets." *Physica A: Statistical Mechanics and its Applications* **388(8)**: pp. 1631 – 1658.

Kristoufek, L. (2012): "How are rescaled range analyses affected by different memory and distributional properties? a monte carlo study." *Physica A: Statistical Mechanics and its Applications* **391**: pp. 4252–4260.

Kuensch, H. R. (1989): "The jackknife and the bootstrap for general stationary observations." *Annals of Statistics* **17**: pp. 1217–1241.

Lo, A. (1991): "Long term memory in stock market prices." *Econometrica* **59**: pp. 1279–1313.

Murphy, A. & M. Izzeldin (2000): "Bootstrapping the small sample critical values of the rescaled range statistic." *The Economic and Social Review* **31**: pp. 351–359.

Murphy, A. & M. Izzeldin (2009): "Bootstrapping long memory tests: Some monte carlo results." *Computational Statistics & Data Analysis* **53(6)**: pp. 2325 – 2334. The Fourth Special Issue on Computational Econometrics.

Palma, W. (2007): *Long-Memory Time Series*. John Wiley & Sons.

Peng, C. K., S. V. Buldyrev, S. Havlin, M. Simons, H. Stanley, & A. Goldberger (1994): "Mosaic organization of dna nucleotides." *Physical Review E* **49**: pp. 1685–1689.

Percival, D. B. & A. T. Walden (2000): *Wavelet Methods for Time Series Analysis (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press.

Robinson, P. M., R. K. Adenstedt, & C. W. J. Granger (2003): *Time Series with Long Memory.* Oxford University Press.

Robinson, P. M. & M. Henry (1999): "Long and short memory conditional heteroskedasticity in estimating the memory parameter of levels." *Econometric theory* **15(3)**: pp. 299–336.

Schmid, F. & M. Trede (2003): "Simple tests for peakedness, fat tails and leptokurtosis based on quantiles." *Comput. Stat. Data Anal.* **43(1)**: pp. 1–12.

Shumway, R. H. & D. S. Stoffer (2006): *Time Series Analysis and Its Applications: With R Examples (Springer Texts in Statistics).* Springer, 2nd edition.

Singh, K. (1981): "On the asymptotic accuracy of efron's bootstrap." *Annals of Statistics* **9**: pp. 1187–1195.

Taylor, S. J. (2000): "Consequences for option pricing of a long memory in volatility." *EFA 2001 Barcelona Meetings, EFMA 2001 Lugano Meetings - working papers* .

Teverovsky, V., M. S. Taqqua, & W. Willinger (1999): "A critical look at lo's modified r/s statistic." *Journal of Statistical Planning and Inference* **80**: pp. 211–227.

Weron, R. (2002): "Estimating long-range dependence: finite sample properties and confidence intervals." *Physica A* **312**: pp. 285 – 299.

Wooldridge, J. (2009): *Introductory Econometrics: A Modern Approach.* ISE - International Student Edition. South Western, Cengage Learning.

Appendix A

Results of Monte Carlo Simulation

Table A.1: Estimated size of tests for WN process

Sample size	Test	Asymptotic Values		Critical	Bootstrapped Values		Critical
		1%	5%	10%	1%	5%	10%
T=512	R/S	0.52	0.59	0.66	0.22	0.33	0.41
	DFA	0.55	0.70	0.75	0.84	0.87	0.90
	GPH	0.00	0.04	0.08	0.34	0.42	0.50
	Wavelet	0.14	0.24	0.33	0.39	0.64	0.73
T=1024	R/S	0.50	0.57	0.63	0.22	0.35	0.42
	DFA	0.56	0.65	0.71	0.79	0.82	0.84
	GPH	0.02	0.06	0.07	0.54	0.61	0.67
	Wavelet	0.11	0.21	0.32	0.70	0.90	0.98

Source: author's computations.

Table A.2: Estimated size of tests for GARCH(1,1) process

Sample size	Test	Asymptotic Values		Critical	Bootstrapped Values		Critical
		1%	5%	10%	1%	5%	10%
T=512	R/S	0.53	0.61	0.66	0.23	0.34	0.42
	DFA	0.50	0.61	0.64	0.83	0.90	0.91
	GPH	0.02	0.06	0.10	0.38	0.44	0.49
	Wavelet	0.10	0.21	0.34	0.39	0.60	0.79
T=1024	R/S	0.54	0.61	0.64	0.22	0.30	0.42
	DFA	0.53	0.65	0.68	0.77	0.82	0.87
	GPH	0.00	0.05	0.08	0.55	0.64	0.68
	Wavelet	0.10	0.16	0.27	0.67	0.90	0.97

Table A.3: Estimated size of tests for ARMA(1,1) process

Sample size	Test	Asymptotic Values		Critical	Bootstrapped Values		Critical
		1%	5%	10%	1%	5%	10%
T=512	R/S	0.87	0.89	0.92	0.63	0.73	0.80
	DFA	0.84	0.94	0.97	1.00	1.00	1.00
	GPH	0.00	0.04	0.11	0.30	0.39	0.49
	Wavelet	0.39	0.57	0.66	0.94	1.00	1.00
T=1024	R/S	0.84	0.90	0.94	0.50	0.57	0.62
	DFA	0.90	0.95	0.95	0.99	0.99	0.99
	GPH	0.02	0.04	0.08	0.54	0.61	0.64
	Wavelet	0.44	0.56	0.65	0.99	1.00	1.00

Table A.4: Estimated power of tests for ARFIMA(0,0.25,0) process

Sample size	Test	Asymptotic Values		Critical	Bootstrapped Values		Critical
		1%	5%	10%	1%	5%	10%
T=512	R/S	0.92	0.95	0.96	0.57	0.67	0.74
	DFA	1.00	1.00	1.00	1.00	1.00	1.00
	GPH	0.15	0.34	0.53	0.30	0.38	0.48
	Wavelet	0.64	0.72	0.75	0.94	0.99	1.00
T=1024	R/S	0.96	0.99	1.00	0.48	0.60	0.66
	DFA	1.00	1.00	1.00	0.99	0.99	0.99
	GPH	0.21	0.46	0.59	0.53	0.62	0.66
	Wavelet	0.84	0.89	0.90	1.00	1.00	1.00

Table A.5: Estimated power of tests for ARFIMA(1,0.25,0) process

Sample size	Test	Asymptotic Values		Critical	Bootstrapped Values		Critical
		1%	5%	10%	1%	5%	10%
T=512	R/S	0.94	0.98	1.00	0.73	0.79	0.86
	DFA	1.00	1.00	1.00	1.00	1.00	1.00
	GPH	0.16	0.38	0.54	0.28	0.38	0.49
	Wavelet	0.72	0.77	0.81	0.99	1.00	1.00
T=1024	R/S	1.00	1.00	1.00	0.59	0.69	0.73
	DFA	1.00	1.00	1.00	0.99	1.00	1.00
	GPH	0.23	0.47	0.58	0.51	0.60	0.67
	Wavelet	0.88	0.90	0.91	1.00	1.00	1.00

Table A.6: Estimated power of tests for ARFIMA(1,-0.25,0) process

Sample size	Test	Asymptotic Values		Critical	Bootstrapped Values		Critical
		1%	5%	10%	1%	5%	10%
T=512	R/S	0.74	0.86	0.87	0.16	0.24	0.29
	DFA	1.00	1.00	1.00	0.24	0.30	0.37
	GPH	0.11	0.22	0.32	0.37	0.47	0.53
	Wavelet	0.88	0.95	0.96	0.03	0.10	0.17
T=1024	R/S	0.87	0.91	0.92	0.38	0.45	0.54
	DFA	1.00	1.00	1.00	0.40	0.54	0.60
	GPH	0.22	0.49	0.60	0.57	0.63	0.69
	Wavelet	0.98	1.00	1.00	0.08	0.20	0.27

Table A.7: Estimated power of tests for ARFIMA(0,0.25,0) with GARCH innovations process

Sample size	Test	Asymptotic Values		Critical	Bootstrapped Values		Critical
		1%	5%	10%	1%	5%	10%
T=512	R/S	1.00	1.00	1.00	0.89	0.94	0.98
	DFA	1.00	1.00	1.00	1.00	1.00	1.00
	GPH	0.95	0.98	0.98	0.72	0.82	0.86
	Wavelet	0.66	0.81	0.83	1.00	1.00	1.00
T=1024	R/S	1.00	1.00	1.00	0.92	0.95	0.97
	DFA	1.00	1.00	1.00	1.00	1.00	1.00
	GPH	0.97	0.99	0.99	0.83	0.91	0.93
	Wavelet	0.94	0.98	1.00	1.00	1.00	1.00

Table A.8: Estimated power of tests for ARFIMA(1,0.25,0) with GARCH innovations process

Sample size	Test	Asymptotic Values	Critical		Bootstrapped Values	Critical	
		1%	5%	10%	1%	5%	10%
T=512	R/S	1.00	1.00	1.00	0.95	0.98	1.00
	DFA	1.00	1.00	1.00	1.00	1.00	1.00
	GPH	0.95	0.97	0.98	0.71	0.82	0.87
	Wavelet	0.80	0.87	0.92	1.00	1.00	1.00
T=1024	R/S	1.00	1.00	1.00	0.93	0.98	0.98
	DFA	1.00	1.00	1.00	1.00	1.00	1.00
	GPH	0.99	1.00	1.00	0.79	0.89	0.93
	Wavelet	0.93	0.95	0.98	1.00	1.00	1.00

Table A.9: Estimated power of tests for ARFIMA(1,-0.25,0) with GARCH innovations process

Sample size	Test	Asymptotic Values	Critical		Bootstrapped Values	Critical	
		1%	5%	10%	1%	5%	10%
T=512	R/S	0.57	0.70	0.76	0.01	0.03	0.07
	DFA	0.92	0.95	0.95	0.08	0.16	0.29
	GPH	0.28	0.42	0.47	0.25	0.39	0.42
	Wavelet	0.82	0.91	0.93	0.15	0.24	0.34
T=1024	R/S	0.69	0.75	0.79	0.11	0.19	0.29
	DFA	0.82	0.89	0.90	0.10	0.21	0.28
	GPH	0.49	0.62	0.66	0.51	0.60	0.64
	Wavelet	0.92	0.97	0.98	0.26	0.48	0.58

Table A.10: SAX and S&P test results

Time Series	Test	Est' of H	St. Error	P-value	Boot' P-value
SAX	R/S	0.689	0.031	0.000	0.090
	DFA	0.506	0.021	0.778	0.260
	GPH	0.800	0.089	0.000	0.010
	Wavelet	0.487	0.023	0.609	0.690
SAX^2	R/S	0.642	0.033	0.000	0.630
	DFA	0.566	0.029	0.021	0.990
	GPH	0.616	0.089	0.190	0.330
	Wavelet	0.538	0.037	0.306	1.000
$\|SAX\|$	R/S	0.718	0.030	0.000	0.800
	DFA	0.637	0.020	0.000	1.000
	GPH	0.769	0.089	0.002	0.010
	Wavelet	0.666	0.021	0.000	1.000
S&P	R/S	0.567	0.026	0.009	0.520
	DFA	0.471	0.015	0.067	0.720
	GPH	0.575	0.061	0.218	0.200
	Wavelet	0.503	0.027	0.897	0.160
$S\&P^2$	R/S	0.682	0.042	0.000	0.650
	DFA	0.681	0.034	0.000	0.970
	GPH	0.804	0.061	0.000	0.010
	Wavelet	0.755	0.060	0.000	0.820
$\|S\&P\|$	R/S	0.809	0.044	0.000	0.650
	DFA	0.777	0.047	0.000	1.000
	GPH	0.968	0.061	0.000	0.740
	Wavelet	0.771	0.058	0.000	1.000

Bachelor Thesis Proposal

Author	Branislav Albert
Supervisor	PhDr. Ladislav Krištoufek
Proposed topic	Long-term memory – detection with bootstrapping techniques

Preliminary thesis content A time series is said to have long-term memory if its current state is in some, possibly non-trivial, way affected by its past performance over a relatively long time span. The problem of its detection using conventional regressions lies inter alia in the fact that they rapidly loose degrees of freedom due to both increasing number of explanatory variables (lagged dependent variables) and decreasing number of observations. There has been proposed a variety of alternative methods to detect the long-term memory, like R/S, GPH and many others. The aim of this paper is to apply moving block bootstrap or some of the other methods to a pseudo-random time series generated with an econometrics program R in order to assess its quality. Processes with long-term memory occur relatively frequently in finance, economics and physics and are thus an ample area of research.

Keywords bootstrapping, moving block bootstrap, long-term memory, time series, R

Klíčová slova

Predbežná náplň práce Časová rada má dlhú pamäť v prípade, že jej súčasný stav je určitým spôsobom ovplyvnený jej priebehom v minulých obdobiach. Jej detekcia klasickými regresnými metódami však naráža na problém nedostatku stupňov voľnosti jednak kvôli rastúcemu počtu premenných (posunuté závislé premenné) a klesajúcemu počtu pozorovaní. Bolo navrhnutých viacero alternatívnych metód detekcie dlhej pamäte, ako R/S, GPH a ďalšie. Cieľom tejto práce je použiť moving block bootstrap alebo nejakú inú metódu na pseudonáhodnú časovú radu vygenerovanú v programe R za účelom zistenia kvality danej metódy. Procesy s dlhou pamäťou sa v ekonómii, financiách a fyzike vyskytujú pomerne často a sú preto vhodnou oblasťou výskumu.

Kľúčové slová bootstrapping, moving block bootstrap, dlhá pamäť, časové rady, R

Outline

1. Introduction
2. Definition of the long-term memory
3. Description of the possible methods
4. Application of the moving block bootstrap on a series with long-term memory
5. Conclusion

Core bibliography

1. Efron, B. (1982): "The jackknife, the bootstrap, and other resampling plans." *Society of Industrial and Applied Mathematics CBMS-NSF Monographs* **38**.
2. Godfrey, L. (2009): "Bootstrap tests for regression models." *Houndmills, GB* : Palgrave Macmillan.
3. Efron, B. & R. J. Tibshirani (1993): "An introduction to the bootstrap." *Boca Raton, FL [US]* : Chapman & Hall.
4. Teyssiere, G. & A. P. Kirman (2007): "Long memory in economics." *DE : Springer* : Heidelberg, DE.
5. Kantelhardt, J. W. (2009): "Encyclopedia of Complexity and Systems Science: Fractal and Multifractal Time Series." *Springer*.
6. Murphy, A. & M. Izzeldin (2008): "Bootstrapping long memory tests: Some Monte Carlo results." *Computational Statistics and Data Analysis* **53**: pp. 2325–2334.
7. Barunik, J. & L. Kristoufek (2008): "On Hurst exponent estimation under heavy-tailed distributions." *Physica A* **389**: pp. 3844–3855.

Author

Supervisor